



B1

<110> Doutriaux, Marie-Pascale  
Betzner, Andreas  
Freyssinet, Georges  
Perez, Pascal

<120> METHOD FOR OBTAINING PLANT VARIETIES

<130> A33153-PCT-USA 072667.0128

<140> US 09/529,239

<141> 2000-10-27

<150> PCT/EP98/06977

<151> 1998-10-09

<160> 98

<210> 1  
<211> 23  
<212> DNA  
<213> Artificial sequence

<220>  
<221> modified\_base  
<222> 11  
<223> I

<220>  
<221> modified\_base  
<222> 14  
<223> I

<220>  
<221> modified\_base  
<222> 17  
<223> I

<220>  
<223> Degenerate oligonucleotides UPMU used to isolate AtMSH3 and AtMSH6.

<300>  
<301> Reenan and Kolodner  
<302> Genetics  
<303> 132  
<306> 963-973  
<307> 1992

<400> 1

ctggatccac nggnccnaay atg 23

<210> 2  
<211> 23  
<212> DNA  
<213> Artificial sequence

<220>

<221> modified\_base  
<222> 15  
<223> I

<220>  
<221> modified\_base  
<222> 18  
<223> I

<220>  
<223> Degenerate oligonucleotides DOMU used to isolate AtMSH3 and AtMSH6.

<300>  
<301> Reenan and Kolodner  
<302> Genetics  
<303> 132  
<306> 963-973  
<307> 1992

<400> 2

ctggatccrt artgngtnrc raa 23

<210> 3  
<211> 24  
<212> DNA  
<213> Artificial sequence

<220>  
<223> MSH3 specific primer 636 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia

<400> 3

tgctagtgcc tcttgcaagc tcat 24

<210> 4  
<211> 27  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Primer AP1 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia containing adapter sequences ligated to both its ends

<400> 4

ccatcctaatacgcactcact atagggc 27

<210> 5  
<211> 23  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Primer AP2 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia containing adapter sequences ligated to both its ends

<400> 5

actcactata gggctcgagc ggc 23

<210> 6

<211> 30

<212> DNA

<213> Artificial sequence

<220>

<223> MSH3 specific primer S525 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia

<400> 6

aggttctgat tatgtgtgac gctttactta 30

<210> 7

<211> 29

<212> DNA

<213> Artificial sequence

<220>

<223> MSH3 specific primer S51 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia

<400> 7

ggatcgggta ctgggttttg agtgtgagg 29

<210> 8

<211> 24

<212> DNA

<213> Artificial sequence

<220>

<223> MSH3 specific primer 635 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia

<400> 8

gcacgtgctt gatggtgttt tcac 24

<210> 9

<211> 28

<212> DNA

<213> Artificial sequence

<220>

<223> MSH3 specific primer S523 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia

<400> 9

tcagacagta tccagcatgg cagaagta 28

<210> 10  
 <211> 33  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> MSH3 specific primer 1S5 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia

<400> 10  
 atccccgggat gggcaagcaa aagcagcaga cga 33

<210> 11  
 <211> 27  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> MSH3 specific primer S53 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia

<400> 11  
 gacaaagagc gaaatgaggc cccttgg 27

<210> 12  
 <211> 1250  
 <212> DNA  
 <213> Arabidopsis thaliana ecotype Columbia  
 <223> Clone 52  
 <400> 12

cccgggatgg	gcaagcaaaa	gcagcagacg	atttctcggt	tcttcgctcc	caaaccctaaa	60
tccccgactc	acgaaccgaa	tccggtagcc	gaatcatcaa	caccgccacc	gaagatatcc	120
gccactgtat	ccttctctcc	ttccaagcgt	aagcttctct	cggaccacct	cgccgcccgcg	180
tcacccaaaa	agcctaaact	ttctcctcac	actcaaaacc	cagtaccoga	tcccaattta	240
caccaaagat	ttctccagag	atctctggaa	ccctcgccgg	aggaatatgt	tcccgaaacg	300
tcattcatga	ggaaatacac	accattggaa	cagcaagtgg	tggagctaaa	gagcaagtac	360
ccagatgtgg	ttttgatggt	ggaagtgggt	tacaggtaca	gattcttcgg	agaagacgcg	420
gagatcgcat	cacgcgtggt	gggtattttac	gctcatatgg	atcacaattt	catgacggcg	480
agtgtgcca	catttcgatt	gaatttccat	gtgagaagac	tggatgaatgc	aggatacaag	540
attgggtgtag	tgaagcagac	tgaaactgca	gccattaagt	cccatgggtgc	aaaccggacc	600
ggcccttttt	tccggggact	gtcggcggtg	tataccaaag	ccacgcttga	agcggctgag	660
gatataagtg	gtggttgtgg	tggatgaagaa	ggttttgggt	cacagagtaa	tttcttgggt	720
tgtgttgtgg	atgagagagt	taagtoggag	acattaggct	gtggatttga	aatgagtttt	780
gatgttagag	tccgtgttgt	tggcggtgaa	atttcgacag	gtgaagtgtg	ttatgaagag	840
ttcaatgata	atttcatgag	aagtggatta	gaggctgtga	ttttgagctt	gtcaccagct	900
gagctgttgc	tgggcccagcc	tctttcacaa	caaactgaga	agtttttgggt	ggcacatgct	960
ggacctacct	caaacgttcg	agtggaaact	gcctcactgg	attgttttcag	caatggtaat	1020
gcagtagatg	aggttatttc	attatgtgaa	aaaatcagcg	caggtaactt	agaagatgat	1080
aaagaaatga	agctggaggc	tgctgaaaaa	ggaatgtctt	gcttgacagt	tcatacaatt	1140
atgaacatgc	cacatctgac	tggtcaagcc	ctcgccctaa	cgtttttgcca	tctcaaacag	1200
tttgatttgg	aaaggatcct	ttaccaaggg	gcctcatttc	gctctttgtc	1250	

<210> 13

<211> 34  
<212> DNA  
<213> Artificial sequence

<220>  
<223> MSH3 specific primer 2S5 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia

<400> 13

atccccgggtc aaaatgaaca agttggtttt agtc 34

<210> 14  
<211> 27  
<212> DNA  
<213> Artificial sequence

<220>  
<223> MSH3 specific primer S52 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia

<400> 14

gccacatctg actgttcaag ccctcgc 27

<210> 15  
<211> 2110  
<212> DNA  
<213> Arabidopsis thaliana ecotype Columbia  
<223> Clone 13

<400> 15

gccacatctg	actgttcaag	ccctcgcct	aacgttttgc	catctcaaac	agtttggatt	60
tgaagagatc	ctttaccaag	gggcctcatt	tcgctctttg	tcaagtaaca	cagagatgac	120
tctctcagcc	aatactctgc	aacagttgga	ggttgtgaaa	aataattcag	atggatcgga	180
atctggctcc	ttattccata	atatgaatca	cacacttaca	gtatatgggt	ccaggcttct	240
tagacactgg	gtgactcatc	ctctatgcga	tagaaatttg	atatctgctc	ggcttgatgc	300
tgtttctgag	atctctgctt	gcatgggagc	tcatagttct	tcccagctca	gcagtgcgtt	360
ggttgaagaa	ggttctgaga	gagcaattgt	atcacctgag	ttttatctcg	tgctctcctc	420
agtcttgaca	gctatgtcta	gatcatctga	tattcaacgt	ggaataacaa	gaatctttca	480
tcggactgct	aaagccacag	agttcattgc	agttatggaa	gctattttac	ttgcggggaa	540
gcaaattcag	cggcttggca	taaagcaaga	ctctgaaatg	aggagtatgc	aatctgcaac	600
tgtgcgatct	actcttttga	gaaaattgat	ttctgttatt	tcacccctg	ttgtggttga	660
caatgccgga	aaactttctc	ctgccctaaa	taagggaagcg	gctgttcgag	gtgacttgct	720
cgacatacta	atcacttcca	gcgaccaatt	tcctgagctt	gctgaagctc	gccaaagcag	780
tttagtcctc	agggaaaagc	tggattcctc	gatagcttca	tttcgcaaga	agctcgctat	840
tcgaaatttg	gaatttcttc	aagtgtcggg	gatcacacat	ttgatagagc	tgcccgttga	900
ttccaaggct	cctatgaatt	gggtgaaagt	aaatagcacc	aagaagacta	ttcgatatca	960
tccccagaaa	atagtagctg	gcttgatgaa	gctagctcta	gcaactgaac	atcttgccat	1020
tgtgaaccga	gcttcgtggg	atagtttcct	caagagtttc	agtagatact	acacagattt	1080
taaggctgcc	gttcaagctc	ttgctgcact	ggactgtttg	cactcccttt	caactctatc	1140
tagaaacaag	aactatgtcc	gtcccagagt	tgtggatgac	tgtgaaccag	ttgagataaa	1200
catacagtc	ggtcgctcat	ctgtactgga	gactatatta	caagataact	tcgtcccaaa	1260
tgacacaatt	ttgcatgcag	aagggggaata	ttgccaaatt	atcacccggc	ctaactggg	1320
aggaaagagc	tgctatatcc	gtcaagtgtc	tttaatttcc	ataatggctc	aggttggttc	1380
ctttgtacca	gcgtcattcg	ccaagctgca	cgtgcttgat	ggtgttttca	ctcggatggg	1440
tgcttcagac	agtatccagc	atggcagaag	tacctttcta	gaagaattaa	gtgaagcgtc	1500

```

acacataatc agaacctggt cttctcggtc gcttggtata ttagatgagc ttggaagagg 1560
cactagcaca caccacggtg tagccattgc ctatgcaaca ttacagcatc tcctagcaga 1620
aaagagatgt ttggttcttt ttgtcacgca ttaccctgaa atagctgaga tcagtaacgg 1680
attcccaggt tctggttgga cataccatgt ctcgatatctg acattgcaga aggataaagg 1740
cagttatgat catgatgatg tgacctacct atataagctt gtgcgtgggtc tttgcagcag 1800
gagctttggt tttaagggtg ctcagcttgc ccagatacct ccatcatgta tacgtcgagc 1860
catttcaatg gctgcaaaat tggaaagctga ggtacgtgca agagagagaa atacacgcat 1920
gggagaacca gaaggacatg aagaaccgag aggcgcagaa gaatctatct cggctctagg 1980
tgacttggtt gcagacctga aatttgctct ctctgaagag gacccttgga aagcattcga 2040
gtttttaaag catgcttgga agattgctgg caaaatcaga ctaaaaccaa cttgttcatt 2100
ttgaccggg 2110

```

```

<210> 16
<211> 29
<212> DNA
<213> Artificial sequence

```

```

<220>
<223> MSH3 specific primer S51 for PCR using cDNA of Arabidopsis thaliana
ecotype Columbia

```

```

<400> 16
ggatcgggta ctgggttttg agtgtgagg 29

```

```

<210> 17
<211> 30
<212> DNA
<213> Artificial sequence

```

```

<220>
<223> MSH3 specific primer S525 for PCR using cDNA of Arabidopsis thaliana
ecotype Columbia

```

```

<400> 17
aggttctgat tatgtgtgac gctttactta 30

```

```

<210> 18
<211> 3522
<212> DNA
<213> Arabidopsis thaliana ecotype Columbia

```

```

<220>
<221> CDS
<222> (100)....(3342)
<223> AtMSH3 full-length cDNA and deduced sequence of the encoded polypeptide

```

```

<400> 18

```

```

cctaagaaag cgcgcgaaaa ttggcaaccc aagttcgcca tagccacgac caccaccttc 60
catttctctt aaacggagga gattacgaat aaagcaatt 99

```

```

atg ggc aag caa aag cag cag acg att tct cgt ttc ttc gct ccc aaa 147
Met Gly Lys Gln Lys Gln Gln Thr Ile Ser Arg Phe Phe Ala Pro Lys
1 5 10 15

```

ccc	aaa	tcc	ccg	act	cac	gaa	ccg	aat	ccg	gta	gcc	gaa	tca	tca	aca	195
Pro	Lys	Ser	Pro	Thr	His	Glu	Pro	Asn	Pro	Val	Ala	Glu	Ser	Ser	Thr	
			20					25					30			
ccg	cca	ccg	aag	ata	tcc	gcc	act	gta	tcc	ttc	tct	cct	tcc	aag	cgt	243
Pro	Pro	Pro	Lys	Ile	Ser	Ala	Thr	Val	Ser	Phe	Ser	Pro	Ser	Lys	Arg	
		35					40					45				
aag	ctt	ctc	tcc	gac	cac	ctc	gcc	gcc	gcg	tca	ccc	aaa	aag	cct	aaa	291
Lys	Leu	Leu	Ser	Asp	His	Leu	Ala	Ala	Ala	Ser	Pro	Lys	Lys	Pro	Lys	
	50					55					60					
ctt	tct	cct	cac	act	caa	aac	cca	gta	ccc	gat	ccc	aat	tta	cac	caa	339
Leu	Ser	Pro	His	Thr	Gln	Asn	Pro	Val	Pro	Asp	Pro	Asn	Leu	His	Gln	
	65				70				75						80	
aga	ttt	ctc	cag	aga	ttt	ctg	gaa	ccc	tcg	ccg	gag	gaa	tat	gtt	ccc	387
Arg	Phe	Leu	Gln	Arg	Phe	Leu	Glu	Pro	Ser	Pro	Glu	Glu	Tyr	Val	Pro	
				85					90					95		
gaa	acg	tca	tca	tcg	agg	aaa	tac	aca	cca	ttg	gaa	cag	caa	gtg	gtg	435
Glu	Thr	Ser	Ser	Ser	Arg	Lys	Tyr	Thr	Pro	Leu	Glu	Gln	Gln	Val	Val	
			100					105					110			
gag	cta	aag	agc	aag	tac	cca	gat	gtg	gtt	ttg	atg	gtg	gaa	gtt	ggc	483
Glu	Leu	Lys	Ser	Lys	Tyr	Pro	Asp	Val	Val	Leu	Met	Val	Glu	Val	Gly	
		115					120					125				
tac	agg	tac	aga	ttc	ttc	gga	gaa	gac	gcg	gag	atc	gca	gca	cgc	gtg	531
Tyr	Arg	Tyr	Arg	Phe	Phe	Gly	Glu	Asp	Ala	Glu	Ile	Ala	Ala	Arg	Val	
	130					135					140					
ttg	ggc	att	tac	gct	cat	atg	gat	cac	aat	ttc	atg	acg	gcg	agt	gtg	579
Leu	Gly	Ile	Tyr	Ala	His	Met	Asp	His	Asn	Phe	Met	Thr	Ala	Ser	Val	
	145				150					155					160	
cca	aca	ttt	cga	ttg	aat	ttc	cat	gtg	aga	aga	ctg	gtg	aat	gca	gga	627
Pro	Thr	Phe	Arg	Leu	Asn	Phe	His	Val	Arg	Arg	Leu	Val	Asn	Ala	Gly	
				165					170					175		
tac	aag	att	ggc	gta	gtg	aag	cag	act	gaa	act	gca	gcc	att	aag	tcc	675
Tyr	Lys	Ile	Gly	Val	Val	Lys	Gln	Thr	Glu	Thr	Ala	Ala	Ile	Lys	Ser	
			180					185					190			
cat	ggc	gca	aac	cgg	acc	ggc	cct	ttt	ttc	cgg	gga	ctg	tcg	gcg	ttg	723
His	Gly	Ala	Asn	Arg	Thr	Gly	Pro	Phe	Phe	Arg	Gly	Leu	Ser	Ala	Leu	
		195					200					205				
tat	acc	aaa	gcc	acg	ctt	gaa	gcg	gct	gag	gat	ata	agt	ggc	ggc	tgt	771
Tyr	Thr	Lys	Ala	Thr	Leu	Glu	Ala	Ala	Glu	Asp	Ile	Ser	Gly	Gly	Cys	
	210					215					220					
ggc	ggc	gaa	gaa	ggc	ttt	ggc	tca	cag	agt	aat	ttc	ttg	gtt	tgt	gtt	819
Gly	Gly	Glu	Glu	Gly	Phe	Gly	Ser	Gln	Ser	Asn	Phe	Leu	Val	Cys	Val	
	225				230					235					240	
gtg	gat	gag	aga	gtt	aag	tcg	gag	aca	tta	ggc	tgt	ggc	att	gaa	atg	867
Val	Asp	Glu	Arg	Val	Lys	Ser	Glu	Thr	Leu	Gly	Cys	Gly	Ile	Glu	Met	
				245					250					255		

agt ttt gat gtt aga gtc ggt gtt gtt ggc gtt gaa att tcg aca ggt	915
Ser Phe Asp Val Arg Val Gly Val Val Gly Val Glu Ile Ser Thr Gly	
260 265 270	
gaa gtt gtt tat gaa gag ttc aat gat aat ttc atg aga agt gga tta	963
Glu Val Val Tyr Glu Glu Phe Asn Asp Asn Phe Met Arg Ser Gly Leu	
275 280 285	
gag gct gtg att ttg agc ttg tca cca gct gag ctg ttg ctt ggc cag	1011
Glu Ala Val Ile Leu Ser Leu Ser Pro Ala Glu Leu Leu Leu Gly Gln	
290 295 300	
cct ctt tca caa caa act gag aag ttt ttg gtg gca cat gct gga cct	1059
Pro Leu Ser Gln Gln Thr Glu Lys Phe Leu Val Ala Met Ala Gly Pro	
305 310 315 320	
acc tca aac gtt cga gtg gaa cgt gcc tca ctg gat tgt ttc agc aat	1107
Thr Ser Asn Val Arg Val Glu Arg Ala Ser Leu Asp Cys Phe Ser Asn	
325 330 335	
ggt aat gca gta gat gag gtt att tca tta tgt gaa aaa atc agc gca	1155
Gly Asn Ala Val Asp Glu Val Ile Ser Leu Cys Glu Lys Ile Ser Ala	
340 345 350	
ggt aac tta gaa gat gat aaa gaa atg aag ctg gag gct gct gaa aaa	1203
Gly Asn Leu Glu Asp Asp Lys Glu Met Lys Leu Glu Ala Ala Glu Lys	
355 360 365	
gga atg tct tgc ttg aca gtt cat aca att atg aac atg cca cat ctg	1251
Gly Met Ser Cys Leu Thr Val His Thr Ile Met Asn Met Pro His Leu	
370 375 380	
act gtt caa gcc ctc gcc cta acg ttt tgc cat ctc aaa cag ttt gga	1299
Thr Val Gln Ala Leu Ala Leu Thr Phe Cys His Leu Lys Gln Phe Gly	
385 390 395 400	
ttt gaa agg atc ctt tac caa ggg gcc tca ttt cgc tct ttg tca agt	1347
Phe Glu Arg Ile Leu Tyr Gln Gly Ala Ser Phe Arg Ser Leu Ser Ser	
405 410 415	
aac aca gag atg act ctc tca gcc aat act ctg caa cag ttg gag gtt	1395
Asn Thr Glu Met Thr Leu Ser Ala Asn Thr Leu Gln Gln Leu Glu Val	
420 425 430	
gtg aaa aat aat tca gat gga tcg gaa tct ggc tcc tta ttc cat aat	1443
Val Lys Asn Asn Ser Asp Gly Ser Glu Ser Gly Ser Leu Phe His Asn	
435 440 445	
atg aat cac aca ctt aca gta tat gct tcc agg ctt ctt aga cac tgg	1491
Met Asn His Thr Leu Thr Val Tyr Gly Ser Arg Leu Leu Arg His Trp	
450 455 460	
gtg act cat cct cta tgc gat aga aat ttg ata tct gct cgg ctt gat	1539
Val Thr His Pro Leu Cys Asp Arg Asn Leu Ile Ser Ala Arg Leu Asp	
465 470 475 480	
gct gtt tct gag att tct gct tgc atg gga tct cat agt tct tcc cag	1587
Ala Val Ser Glu Ile Ser Ala Cys Met Gly Ser His Ser Ser Ser Gln	
485 490 495	



ctc agc agt gag ttg gtt gaa gaa ggt tct gag aga gca att gta tca	1635
Leu Ser Ser Glu Leu Val Glu Glu Gly Ser Glu Arg Ala Ile Val Ser	
500 505 510	
cct gag ttt tat ctc gtg ctc tcc tca gtc ttg aca gct atg tct aga	1683
Pro Glu Phe Tyr Leu Val Leu Ser Ser Val Leu Thr Ala Met Ser Arg	
515 520 525	
tca tct gat att caa cgt gga ata aca aga atc ttt cat cgg act gct	1731
Ser Ser Asp Ile Gln Arg Gly Ile Thr Arg Ile Phe His Arg Thr Ala	
530 535 540	
aaa gcc aca gag ttc att gca gtt atg gaa gct att tta ctt gcg ggg	1779
Lys Ala Thr Glu Phe Ile Ala Val Met Glu Ala Ile Leu Leu Ala Gly	
545 550 555 560	
aag caa att cag cgg ctt ggc ata aag caa gac tct gaa atg agg agt	1827
Lys Gln Ile Gln Arg Leu Gly Ile Lys Gln Asp Ser Glu Met Arg Ser	
565 570 575	
atg caa tct gca act gtg cga tct act ctt ttg aga aaa ttg att tct	1875
Met Gln Ser Ala Thr Val Arg Ser Thr Leu Leu Arg Lys Leu Ile Ser	
580 585 590	
gtt att tca tcc cct gtt gtg gtt gac aat gcc gga aaa ctt ctc tct	1923
Val Ile Ser Ser Pro Val Val Val Asp Asn Ala Gly Lys Leu Leu Ser	
595 600 605	
gcc cta aat aag gaa gcg gct gtt cga ggt gac ttg ctc gac ata cta	1971
Ala Leu Asn Lys Glu Ala Ala Val Arg Gly Asp Leu Leu Asp Ile Leu	
610 615 620	
atc act tcc agc gac caa ttt cct gag ctt gct gaa gct cgc caa gca	2019
Ile Thr Ser Ser Asp Gln Phe Pro Glu Leu Ala Glu Ala Arg Gln Ala	
625 630 635 640	
gtt tta gtc atc agg gaa aag ctg gat tcc tcg ata gct tca ttt cgc	2067
Val Leu Val Ile Arg Glu Lys Leu Asp Ser Ser Ile Ala Ser Phe Arg	
645 650 655	
aag aag ctc gct att cga aat ttg gaa ttt ctt caa gtg tcg ggg atc	2115
Lys Lys Leu Ala Ile Arg Asn Leu Glu Phe Leu Gln Val Ser Gly Ile	
660 665 670	
aca cat ttg ata gag ctg ccc gtt gat tcc aag gtc cct atg aat tgg	2163
Thr His Leu Ile Glu Leu Pro Val Asp Ser Lys Val Pro His Asn Trp	
675 680 685	
gtg aaa gta aat agc acc aag aag act att cga tat cat ccc cca gaa	2211
Val Lys Val Asn Ser Thr Lys Lys Thr Ile Arg Tyr His Pro Pro Glu	
690 695 700	
ata gta gct ggc ttg gat gag cta gct cta gca act gaa cat ctt gcc	2259
Ile Val Ala Gly Leu Asp Glu Leu Ala Leu Ala Thr Glu His Leu Ala	
705 710 715 720	
att gtg aac cga gct tcg tgg gat agt ttc ctc aag agt ttc agt aga	2307
Ile Val Asn Arg Ala Ser Trp Asp Ser Phe Leu Lys Ser Phe Ser Arg	
725 730 735	

tac tac aca gat ttt aag gct gcc gtt caa gct ctt gct gca ctg gac	2355
Tyr Tyr Thr Asp Phe Lys Ala Ala Val Gln Ala Leu Ala Ala Leu Asp	
740 745 750	
tgt ttg cac tcc ctt tca act cta tct aga aac aag aac tat gtc cgt	2403
Cys Leu His Ser Leu Ser Thr Leu Ser Arg Asn Lys Asn Tyr Val Arg	
755 760 765	
ccc gag ttt gtg gat gac tgt gaa cca gtt gag ata aac ata cag tct	2451
Pro Glu Phe Val Asp Asp Cys Glu Pro Val Glu Ile Asn Ile Gln Ser	
770 775 780	
ggc cgt cat cct gta ctg gag act ata tta caa gat aac ttc gtc cca	2499
Gly Arg His Pro Val Leu Glu Thr Ile Leu Gln Asp Asn Phe Val Pro	
785 790 795 800	
aat gac aca att ttg cat gca gaa ggg gaa tat tgc caa att atc acc	2547
Asn Asp Thr Ile Leu His Ala Glu Gly Glu Tyr Cys Gln Ile Ile Thr	
805 810 815	
gga cct aac atg gga gga aag agc tgc tat atc cgt caa gtt gct tta	2595
Gly Pro Asn Met Gly Gly Lys Ser Cys Tyr Ile Arg Gln Val Ala Leu	
820 825 830	
att tcc ata atg gct cag gtt ggt tcc ttt gta cca gcg tca ttc gcc	2643
Ile Ser Ile Met Ala Gln Val Gly Ser Phe Val Pro Ala Ser Phe Ala	
835 840 845	
aag ctg cac gtg ctt gat ggt gtt ttc act cgg atg ggt gct tca gac	2691
Lys Leu His Val Leu Asp Gly Val Phe Thr Arg Met Gly Ala Ser Asp	
850 855 860	
agt atc cag cat ggc aga agt acc ttt cta gaa gaa tta agt gaa gcg	2739
Ser Ile Gln His Gly Arg Ser Thr Phe Leu Glu Glu Leu Ser Glu Ala	
865 870 875 880	
tca cac ata atc aga acc tgt tct tct cgt tcg ctt gtt ata tta gat	2787
Ser His Ile Ile Arg Thr Cys Ser Ser Arg Ser Leu Val Ile Leu Asp	
885 890 895	
gag ctt gga aga ggc act agc aca cac gac ggt gta gcc att gcc tat	2835
Glu Leu Gly Arg Gly Thr Ser Thr His Asp Gly Val Ala Ile Ala Tyr	
900 905 910	
gca aca tta cag cat ctc cta gca gaa aag aga tgt ttg gtt ctt ttt	2883
Ala Thr Leu Gln His Leu Leu Ala Glu Lys Arg Cys Leu Val Leu Phe	
915 920 925	
gtc acg cat tac cct gaa ata gct gag atc agt aac gga ttc cca ggt	2931
Val Thr His Tyr Pro Glu Ile Ala Glu Ile Ser Asn Gly Phe Pro Gly	
930 935 940	
tct gtt ggg aca tac cat gtc tcg tat ctg aca ttg cag aag gat aaa	2979
Ser Val Gly Thr Tyr His Val Ser Tyr Leu Thr Leu Gln Lys Asp Lys	
945 950 955 960	
ggc agt tat gat cat gat gat gtg acc tac cta tat aag ctt gtg cgt	3027
Gly Ser Tyr Asp His Asp Asp Val Thr Tyr Leu Tyr Lys Leu Val Arg	
965 970 975	

ggt ctt tgc agc agg agc ttt ggt ttt aag gtt gct cag ctt gcc cag 3075  
 Gly Leu Cys Ser Arg Ser Phe Gly Phe Lys Val Ala Gln Leu Ala Gln  
 980 985 990  
 ata cct cca tca tgt ata cgt cga gcc att tca atg gct gca aaa ttg 3123  
 Ile Pro Pro Ser Cys Ile Arg Arg Ala Ile Ser Met Ala Ala Lys Leu  
 995 1000 1005  
 gaa gct gag gta cgt gca aga gag aga aat aca cgc atg gga gaa cca 3171  
 Glu Ala Glu Val Arg Ala Arg Glu Arg Asn Thr Arg Met Gly Glu Pro  
 1010 1015 1020  
 gaa gga cat gaa gaa ccg aga ggc gca gaa gaa tct att tcg gct cta 3219  
 Glu Gly His Glu Glu Pro Arg Gly Ala Glu Glu Ser Ile Ser Ala Leu  
 1025 1030 1035 1040  
 ggt gac ttg ttt gca gac ctg aaa ttt gct ctc tct gaa gag gac cct 3267  
 Gly Asp Leu Phe Ala Asp Leu Lys Phe Ala Leu Ser Glu Glu Asp Pro  
 1045 1050 1055  
 tgg aaa gca ttc gag ttt tta aag cat gct tgg aag att gct ggc aaa 3315  
 Trp Lys Ala Phe Glu Phe Leu Lys His Ala Trp Lys Ile Ala Gly Lys  
 1060 1065 1070  
 atc aga cta aaa cca act tgt tca ttt tgatttaatc ttaacattat 3362  
 Ile Arg Leu Lys Pro Thr Cys Ser Phe  
 1075 1080  
 agcaactgca aggtcttgat catctgtag ttgcgtacta acttatgtgt attagtataa 3422  
 caagaaaaga gaattagaga gatggattct aatccgggtgt tgcagtacat cttttctcca 3482  
 cccgcataaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3522

<210> 19  
 <211> 1081  
 <212> PRT  
 <213> Arabidopsis thaliana ecotype Columbia  
 <223> Polypeptide MSH3

<400> 19

Met Gly Lys Gln Lys Gln Gln Thr Ile Ser Arg Phe Phe Ala Pro Lys  
 1 5 10 15  
 Pro Lys Ser Pro Thr His Glu Pro Asn Pro Val Ala Glu Ser Ser Thr  
 20 25 30  
 Pro Pro Pro Lys Ile Ser Ala Thr Val Ser Phe Ser Pro Ser Lys Arg  
 35 40 45  
 Lys Leu Leu Ser Asp His Leu Ala Ala Ala Ser Pro Lys Lys Pro Lys  
 50 55 60  
 Leu Ser Pro His Thr Gln Asn Pro Val Pro Asp Pro Asn Leu His Gln  
 65 70 75 80  
 Arg Phe Leu Gln Arg Phe Leu Glu Pro Ser Pro Glu Glu Tyr Val Pro  
 85 90 95  
 Glu Thr Ser Ser Ser Arg Lys Tyr Thr Pro Leu Glu Gln Gln Val Val

100					105					110						
Glu	Leu	Lys	Ser	Lys	Tyr	Pro	Asp	Val	Val	Leu	Met	Val	Glu	Val	Gly	
115					120					125						
Tyr	Arg	Tyr	Arg	Phe	Phe	Gly	Glu	Asp	Ala	Glu	Ile	Ala	Ala	Arg	Val	
130					135					140						
Leu	Gly	Ile	Tyr	Ala	His	Met	Asp	His	Asn	Phe	Met	Thr	Ala	Ser	Val	
145					150					155					160	
Pro	Thr	Phe	Arg	Leu	Asn	Phe	His	Val	Arg	Arg	Leu	Val	Asn	Ala	Gly	
165					170					175						
Tyr	Lys	Ile	Gly	Val	Val	Lys	Gln	Thr	Glu	Thr	Ala	Ala	Ile	Lys	Ser	
180					185					190						
His	Gly	Ala	Asn	Arg	Thr	Gly	Pro	Phe	Phe	Arg	Gly	Leu	Ser	Ala	Leu	
195					200					205						
Tyr	Thr	Lys	Ala	Thr	Leu	Glu	Ala	Ala	Glu	Asp	Ile	Ser	Gly	Gly	Cys	
210					215					220						
Gly	Gly	Glu	Glu	Gly	Phe	Gly	Ser	Gln	Ser	Asn	Phe	Leu	Val	Cys	Val	
225					230					235					240	
Val	Asp	Glu	Arg	Val	Lys	Ser	Glu	Thr	Leu	Gly	Cys	Gly	Ile	Glu	Met	
245					250					255						
Ser	Phe	Asp	Val	Arg	Val	Gly	Val	Val	Gly	Val	Glu	Ile	Ser	Thr	Gly	
260					265					270						
Glu	Val	Val	Tyr	Glu	Glu	Phe	Asn	Asp	Asn	Phe	Met	Arg	Ser	Gly	Leu	
275					280					285						
Glu	Ala	Val	Ile	Leu	Ser	Leu	Ser	Pro	Ala	Glu	Leu	Leu	Leu	Gly	Gln	
290					295					300						
Pro	Leu	Ser	Gln	Gln	Thr	Glu	Lys	Phe	Leu	Val	Ala	Met	Ala	Gly	Pro	
305					310					315					320	
Thr	Ser	Asn	Val	Arg	Val	Glu	Arg	Ala	Ser	Leu	Asp	Cys	Phe	Ser	Asn	
325					330					335						
Gly	Asn	Ala	Val	Asp	Glu	Val	Ile	Ser	Leu	Cys	Glu	Lys	Ile	Ser	Ala	
340					345					350						
Gly	Asn	Leu	Glu	Asp	Asp	Lys	Glu	Met	Lys	Leu	Glu	Ala	Ala	Glu	Lys	
355					360					365						
Gly	Met	Ser	Cys	Leu	Thr	Val	His	Thr	Ile	Met	Asn	Met	Pro	His	Leu	
370					375					380						
Thr	Val	Gln	Ala	Leu	Ala	Leu	Thr	Phe	Cys	His	Leu	Lys	Gln	Phe	Gly	
385					390					395					400	
Phe	Glu	Arg	Ile	Leu	Tyr	Gln	Gly	Ala	Ser	Phe	Arg	Ser	Leu	Ser	Ser	
405					410					415						
Asn	Thr	Glu	Met	Thr	Leu	Ser	Ala	Asn	Thr	Leu	Gln	Gln	Leu	Glu	Val	

420					425					430					
Val	Lys	Asn	Asn	Ser	Asp	Gly	Ser	Glu	Ser	Gly	Ser	Leu	Phe	His	Asn
		435					440					445			
Met	Asn	His	Thr	Leu	Thr	Val	Tyr	Gly	Ser	Arg	Leu	Leu	Arg	His	Trp
	450					455					460				
Val	Thr	His	Pro	Leu	Cys	Asp	Arg	Asn	Leu	Ile	Ser	Ala	Arg	Leu	Asp
465					470					475					480
Ala	Val	Ser	Glu	Ile	Ser	Ala	Cys	Met	Gly	Ser	His	Ser	Ser	Ser	Gln
				485					490						495
Leu	Ser	Ser	Glu	Leu	Val	Glu	Glu	Gly	Ser	Glu	Arg	Ala	Ile	Val	Ser
			500					505					510		
Pro	Glu	Phe	Tyr	Leu	Val	Leu	Ser	Ser	Val	Leu	Thr	Ala	Met	Ser	Arg
		515					520					525			
Ser	Ser	Asp	Ile	Gln	Arg	Gly	Ile	Thr	Arg	Ile	Phe	His	Arg	Thr	Ala
	530					535					540				
Lys	Ala	Thr	Glu	Phe	Ile	Ala	Val	Met	Glu	Ala	Ile	Leu	Leu	Ala	Gly
545					550					555					560
Lys	Gln	Ile	Gln	Arg	Leu	Gly	Ile	Lys	Gln	Asp	Ser	Glu	Met	Arg	Ser
				565					570					575	
Met	Gln	Ser	Ala	Thr	Val	Arg	Ser	Thr	Leu	Leu	Arg	Lys	Leu	Ile	Ser
			580					585					590		
Val	Ile	Ser	Ser	Pro	Val	Val	Val	Asp	Asn	Ala	Gly	Lys	Leu	Leu	Ser
		595					600					605			
Ala	Leu	Asn	Lys	Glu	Ala	Ala	Val	Arg	Gly	Asp	Leu	Leu	Asp	Ile	Leu
	610					615					620				
Ile	Thr	Ser	Ser	Asp	Gln	Phe	Pro	Glu	Leu	Ala	Glu	Ala	Arg	Gln	Ala
625					630					635					640
Val	Leu	Val	Ile	Arg	Glu	Lys	Leu	Asp	Ser	Ser	Ile	Ala	Ser	Phe	Arg
				645					650					655	
Lys	Lys	Leu	Ala	Ile	Arg	Asn	Leu	Glu	Phe	Leu	Gln	Val	Ser	Gly	Ile
		660						665					670		
Thr	His	Leu	Ile	Glu	Leu	Pro	Val	Asp	Ser	Lys	Val	Pro	His	Asn	Trp
		675					680					685			
Val	Lys	Val	Asn	Ser	Thr	Lys	Lys	Thr	Ile	Arg	Tyr	His	Pro	Pro	Glu
	690					695					700				
Ile	Val	Ala	Gly	Leu	Asp	Glu	Leu	Ala	Leu	Ala	Thr	Glu	His	Leu	Ala
705					710					715					720
Ile	Val	Asn	Arg	Ala	Ser	Trp	Asp	Ser	Phe	Leu	Lys	Ser	Phe	Ser	Arg
				725					730					735	
Tyr	Tyr	Thr	Asp	Phe	Lys	Ala	Ala	Val	Gln	Ala	Leu	Ala	Ala	Leu	Asp

740					745					750					
Cys	Leu	His	Ser	Leu	Ser	Thr	Leu	Ser	Arg	Asn	Lys	Asn	Tyr	Val	Arg
	755						760					765			
Pro	Glu	Phe	Val	Asp	Asp	Cys	Glu	Pro	Val	Glu	Ile	Asn	Ile	Gln	Ser
	770					775					780				
Gly	Arg	His	Pro	Val	Leu	Glu	Thr	Ile	Leu	Gln	Asp	Asn	Phe	Val	Pro
	785					790					795				800
Asn	Asp	Thr	Ile	Leu	His	Ala	Glu	Gly	Glu	Tyr	Cys	Gln	Ile	Ile	Thr
				805							810				815
Gly	Pro	Asn	Met	Gly	Gly	Lys	Ser	Cys	Tyr	Ile	Arg	Gln	Val	Ala	Leu
			820					825					830		
Ile	Ser	Ile	Met	Ala	Gln	Val	Gly	Ser	Phe	Val	Pro	Ala	Ser	Phe	Ala
			835					840					845		
Lys	Leu	His	Val	Leu	Asp	Gly	Val	Phe	Thr	Arg	Met	Gly	Ala	Ser	Asp
	850					855					860				
Ser	Ile	Gln	His	Gly	Arg	Ser	Thr	Phe	Leu	Glu	Glu	Leu	Ser	Glu	Ala
	865					870					875				880
Ser	His	Ile	Ile	Arg	Thr	Cys	Ser	Ser	Arg	Ser	Leu	Val	Ile	Leu	Asp
				885					890						895
Glu	Leu	Gly	Arg	Gly	Thr	Ser	Thr	His	Asp	Gly	Val	Ala	Ile	Ala	Tyr
			900					905						910	
Ala	Thr	Leu	Gln	His	Leu	Leu	Ala	Glu	Lys	Arg	Cys	Leu	Val	Leu	Phe
			915					920					925		
Val	Thr	His	Tyr	Pro	Glu	Ile	Ala	Glu	Ile	Ser	Asn	Gly	Phe	Pro	Gly
	930					935						940			
Ser	Val	Gly	Thr	Tyr	His	Val	Ser	Tyr	Leu	Thr	Leu	Gln	Lys	Asp	Lys
	945					950					955				960
Gly	Ser	Tyr	Asp	His	Asp	Asp	Val	Thr	Tyr	Leu	Tyr	Lys	Leu	Val	Arg
				965					970					975	
Gly	Leu	Cys	Ser	Arg	Ser	Phe	Gly	Phe	Lys	Val	Ala	Gln	Leu	Ala	Gln
			980					985					990		
Ile	Pro	Pro	Ser	Cys	Ile	Arg	Arg	Ala	Ile	Ser	Met	Ala	Ala	Lys	Leu
			995					1000					1005		
Glu	Ala	Glu	Val	Arg	Ala	Arg	Glu	Arg	Asn	Thr	Arg	Met	Gly	Glu	Pro
	1010							1015					1020		
Glu	Gly	His	Glu	Glu	Pro	Arg	Gly	Ala	Glu	Glu	Ser	Ile	Ser	Ala	Leu
	1025					1030					1035				1040
Gly	Asp	Leu	Phe	Ala	Asp	Leu	Lys	Phe	Ala	Leu	Ser	Glu	Glu	Asp	Pro
			1045						1050					1055	
Trp	Lys	Ala	Phe	Glu	Phe	Leu	Lys	His	Ala	Trp	Lys	Ile	Ala	Gly	Lys

	1060	1065	1070
Ile Arg Leu Lys Pro Thr Cys Ser Phe			
	1075	1080	

<210> 20  
 <211> 24  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> MSH6 specific primer 638 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia

<400> 20  
 tctctaccag gtgacgaaaa accg 24

<210> 21  
 <211> 28  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> Primer S81 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia

<400> 21  
 cgtcgccttt agcateccct tccttcac 28

<210> 22  
 <211> 30  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> MSH6 specific primer S823 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia

<400> 22  
 gcttggcgca tctaataagaa tcatgacagg 30

<210> 23  
 <211> 24  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> MSH6 specific primer 637 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia

<400> 23  
 gacagcgtca gttcttcaga atgc 24

<210> 24  
 <211> 33  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> MSH6 specific primer 1S8 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia

<400> 24  
 atccccgggat gcagcgccag agatcgattt tgt 33

<210> 25  
 <211> 27  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> MSH6 specific primer S83 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia

<400> 25  
 cgctatctat ggctgcttcg aattgag 27

<210> 26  
 <211> 1385  
 <212> DNA  
 <213> Arabidopsis thaliana ecotype Columbia  
 <223> Clone 43  
 <400> 26

ccccgggatgc	agcgccagag	atcgattttg	tctttcttcc	aaaaaccac	ggcggcgact	60
acgaagggtt	tggtttccgg	cgatgctgct	agcggcgggg	gcggcagcgg	aggaccacga	120
tttaatgtga	aggaaggga	tgctaaaggc	gacgcttctg	tacgttttgc	tgtttcgaaa	180
tctgtcgatg	aggtagagg	aacggatact	ccaccggaga	aggttccgcg	tcgtgtcctg	240
ccgtctggat	ttaagccggc	tgaatccgcc	ggtgatgctt	cgtccctggt	ctccaatatt	300
atgcataagt	ttgtaaaagt	cgatgatcga	gattgttctg	gagagaggag	ccgagaagat	360
gttgttccgc	tgaatgattc	atctctatgt	atgaaggcta	atgatgttat	tcctcaattt	420
cgttccaata	atggtaaaac	tcaagaaaga	aaccatgctt	ttagtttcag	tgggagagct	480
gaacttagat	cagtagaaga	tataggagta	gatggcgatg	ttcctgggtc	agaaacacca	540
gggatgcgc	cacgtgcttc	tcgcttgaag	cgagtcttgg	aggatgaaat	gacttttaag	600
gaggataagg	ttcctgtatt	ggactctaac	aaaaggctga	aaatgctcca	ggatccgggt	660
tgtggagaga	agaaagaagt	aaacgaagga	accaaatttg	aatggcttga	gtcttctcga	720
atcaggggatg	ccaatagaag	acgtcctgat	gatccccttt	acgatagaaa	gaccttacac	780
ataccacctg	atgttttcaa	gaaaatgtct	gcatcacaaa	agcaatattg	gagtgttaag	840
agtgaatata	tggacattgt	gcttttcttt	aaagtgggga	aattttatga	gctgtatgag	900
ctagatgcgg	aattaggtca	caaggagctt	gactggaaga	tgaccatgag	tgggtgggga	960
aaatgcagac	aggttggtat	ctctgaaagt	gggatagatg	aggcagtgca	aaagctatta	1020
gctcgtggat	ataaagttgg	acgaatcgag	cagctagaaa	catctgacca	agcaaaaagcc	1080
agaggtgcta	atactataat	tccaaggaag	ctagttcagg	tattaactcc	atcaacagca	1140
agcgaaggaa	acatcgggcc	tgatgccgtc	catcttcttg	ctataaaaaga	gatcaaaatg	1200
gagctacaaa	agtgttcaac	tgtgtatgga	tttgcttttg	ttgactgtgc	tgcttgagg	1260
ttttgggttg	ggtccatcag	cgatgatgca	tcatgtgctg	ctcttgagag	gttattgatg	1320
caggtttctc	caaaggaagt	gttatatgac	agtaaaaggg	tatcaagaga	agcacaaaag	1380
gctctaagga	aatatacggt	gacaggggtc	acggcgggtac	agttgggtcc	agtaccacaa	1440



gtaatggggg	atacagatgc	tgctggagtt	agaaatataa	tagaatctaa	cggatacttt	1500
aaaggtttct	ctgaatcatg	gaactgtgct	gttgatgggc	taaataaatg	tgatgttgcc	1560
cttagtgctc	ttggagagct	aattaatcat	ctgtctaggc	taaagctaga	agatgtactt	1620
aagcatgggg	atatttttcc	ataccaagtt	tacagggggt	gtctcagaat	tgatggccag	1680
acgatggtaa	atcttgagat	atttaacaat	agctgtgatg	gtggtccttc	agggaccttg	1740
tacaaatatc	ttgataactg	tgttagtcca	actggtaagc	gactcttaag	gaattggatc	1800
tgccatccac	tcaaagatgt	agaaagcatc	aataaacggc	ttgatgtagt	tgaagaattc	1860
acggcaaaact	cagaaagtat	gcaaatacact	ggccagtatc	tccacaaaact	tccagactta	1920
gaaagactgc	tccgacgcac	caagtctagc	gttcgatcat	cagcctctgt	gttgccctgct	1980
cttctgggga	aaaaagtgtc	gaaacaacga	gttaaagcat	ttgggcaaat	tgtgaaaggg	2040
ttcagaagtg	gaattgatct	gttggttggt	ctacagaagg	aatcaaatat	gatgagtgtt	2100
ctttataaac	tcgtgaaact	tcctatatatta	gtaggaaaaa	gcgggctaga	gttatttctt	2160
tctcaattcg	aagcagccat	agatagcg			2188	

<210> 27  
 <211> 1385  
 <212> DNA  
 <213> Arabidopsis thaliana ecotype Columbia  
 <223> Clone 62

<400> 27

catcagcctc	tgtgttgctc	gctcttctgg	ggaaaaaagt	gctgaaacaa	cagagttaaag	60
catttgggca	aattgtgaaa	gggttcagaa	gtggaattga	tctgttgttg	gctctacaga	120
aggaatcaaa	tatgatgagt	ttgctttata	aactctgtaa	acttctctata	ttagtaggaa	180
aaagcgggct	agagttatct	ctttctcaat	tcgaagcagc	catagatagc	gactttccaa	240
attatcagaa	ccaagatgtg	acagatgaaa	acgctgaaaac	tctcacaata	cttatcgaac	300
tttttatcga	aagagcaact	caatggtctg	aggtcattca	caccataaagc	tcgcttagatg	360
tcctgagcac	ttttgcaatc	gcagcaagtc	tctctgctgg	aagcatggcc	aggcctgtta	420
tttttcccga	atcagaagct	acagatcaga	atcagaaaaac	aaaaggggcca	atacttaaaa	480
tccaaggact	atggcatcca	tttgcatgtg	cagccgatgg	tcaattgcct	gttccgaatg	540
atatactcct	tggcgaggct	agaagaagca	gtggcagcat	tcctcctcgg	tcattgttac	600
tgacgggacc	aaacatgggc	ggaaaatcaa	ctcttctctg	tgcaacatgt	ctggccgtta	660
tctttgcccc	acttggtctg	tacgtgccgt	gtgagtcctt	cgaaatctcc	ctcgtggata	720
ctatcttcac	aaggcttggc	gcatctgata	gaatcatgac	aggagagagt	acctttttgg	780
tagaatgcac	tgagacagcg	tcagttcttc	agaatgcaac	tcaggattca	ctagtaattc	840
ttgacgaact	gggcagagga	actagtactt	tcgatggata	cgccattgca	tactcggttt	900
ttcgtcacct	ggtagagaaa	gttcaatgtc	ggatgctctt	tgcaacacat	taccaccctc	960
tcaccaagga	attcgcgtct	caccacgtg	tcacctcgaa	acacatggct	tgcgcatcca	1020
aatcaagatc	tgattatcaa	ccacgtgggt	gtgatcaaga	cctagtgttc	ttgtaccggt	1080
taaccgaggg	agcttgtcct	gagagctacg	gacttcaagt	ggcactcatg	gctggaatac	1140
caaaaccaagt	ggttgaaaca	gcatcaggtg	ctgctcaagc	catgaagaga	tcaattgggg	1200
aaaacttcaa	gtcaagttag	ctaagatctg	agttctcaag	tctgcatgaa	gactggctca	1260
agtcattggg	gggtatttct	cgagtcgccc	acaacaatgc	ccccattggc	gaagatgact	1320
acgacacttt	gttttgctta	tggcatgaga	tcaaatactc	ttactgtgtt	cccaaataac	1380
ccggg					1385	

<210> 28  
 <211> 34  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> MSH6 specific primer 2S8 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia

<400> 26

atccccgggtt atttggaac acagtaagag gatt 34

<210> 29  
<211> 27  
<212> DNA  
<213> Artificial sequence

<220>  
<223> MSH6 specific primer S82 for PCR using cDNA of Arabidopsis thaliana ecotype Columbia

<400> 29

gcgttcgatc atcagcctct gtgttgc 27

<210> 30  
<211> 3606  
<212> DNA  
<213> Arabidopsis thaliana ecotype Columbia

<220>  
<221> CDS  
<222> (142)....(3468)  
<223> AtMSH6 full-length cDNA and deduced sequence of the encoded polypeptide

<400> 30

aaaagttgag ccctgaggag tatcgtttcc gccatttcta cgacgcaagg cgaaaatttt 60  
tggcgccaat ctttcccccc tttcgaattc tctcagctca aaacatcgtt tctctctcac 120

tctctctcac aattccaaaa a atg cag cgc cag aga tcg att ttg tct ttc 171  
Met Gln Arg Gln Arg Ser Ile Leu Ser Phe  
1 5 10

ttc caa aaa ccc acc gcg gcg act acg aag ggt ttg gtt tcc ggc gat 219  
Phe Gln Lys Pro Thr Ala Ala Thr Thr Lys Gly Leu Val Ser Gly Asp  
15 20 25

gct gct agc ggc ggg ggc ggc agc gga gga cca cga ttt aat gtg aag 267  
Ala Ala Ser Gly Gly Gly Gly Ser Gly Gly Pro Arg Phe Asn Val Arg  
30 35 40

gaa ggg gat gct aaa ggc gac gct tct gta cgt ttt gct gtt tcg aaa 315  
Glu Gly Asp Ala Lys Gly Asp Ala Ser Val Arg Phe Ala Val Ser Lys  
45 50 55

tct gtc gat gag gtt aga gga acg gat act cca ccg gag aag gtt ccg 363  
Ser Val Asp Glu Val Arg Gly Thr Asp Thr Pro Pro Glu Lys Val Pro  
60 65 70

cgt cgt gtc ctg ccg tct gga ttt aag ccg gct gaa tcc gcc gst gat 411  
Arg Arg Val Leu Pro Ser Gly Phe Lys Pro Ala Glu Ser Ala Gly Asp  
75 80 85 90

gct tcg tcc ctg ttc tcc aat att atg cat aag ttt gta aaa gtc gat 459  
Ala Ser Ser Leu Phe Ser Asn Ile Met His Lys Phe Val Lys Val Asp  
95 100 105

gat	cga	gat	tgt	tct	gga	gag	agg	agc	cga	gaa	gat	ggt	ggt	ccg	ctg	507
Asp	Arg	Asp	Cys	Ser	Gly	Glu	Arg	Ser	Arg	Glu	Asp	Val	Val	Pro	Leu	
			110					115					120			
aat	gat	tca	tct	cta	tgt	atg	aag	gct	aat	gat	ggt	att	cct	caa	ttt	555
Asn	Asp	Ser	Ser	Leu	Cys	Met	Lys	Ala	Asn	Asp	Val	Ile	Pro	Gln	Phe	
			125				130					135				
cgt	tcc	aat	aat	ggg	aaa	act	caa	gaa	aga	aac	cat	gct	ttt	agt	ttc	603
Arg	Ser	Asn	Asn	Gly	Lys	Thr	Gln	Glu	Arg	Asn	His	Ala	Phe	Ser	Phe	
			140				145				150					
agt	ggg	aga	gct	gaa	ctt	aga	tca	gta	gaa	gat	ata	gga	gta	gat	ggc	651
Ser	Gly	Arg	Ala	Glu	Leu	Arg	Ser	Val	Glu	Asp	Ile	Gly	Val	Asp	Gly	
					160					165					170	
gat	ggt	cct	ggg	cca	gaa	aca	cca	ggg	atg	cgt	cca	cgt	gct	tct	cgc	699
Asp	Val	Pro	Gly	Pro	Glu	Thr	Pro	Gly	Met	Arg	Pro	Arg	Ala	Ser	Arg	
				175					180					185		
ttg	aag	cga	ggt	ctg	gag	gat	gaa	atg	act	ttt	aag	gag	gat	aag	ggt	747
Leu	Lys	Arg	Val	Leu	Glu	Asp	Glu	Met	Thr	Phe	Lys	Glu	Asp	Lys	Val	
			190					195					200			
cct	gta	ttg	gac	tct	aac	aaa	agg	ctg	aaa	atg	ctc	cag	gat	ccg	ggt	795
Pro	Val	Leu	Asp	Ser	Asn	Lys	Arg	Leu	Lys	Met	Leu	Gln	Asp	Pro	Val	
			205				210					215				
tgt	gga	gag	aag	aaa	gaa	gta	aac	gaa	gga	acc	aaa	ttt	gaa	tgg	ctt	843
Cys	Gly	Glu	Lys	Lys	Glu	Val	Asn	Glu	Gly	Thr	Lys	Phe	Glu	Trp	Leu	
			220				225				230					
gag	tct	tct	cga	atc	agg	gat	gcc	aat	aga	aga	cgt	cct	gat	gat	ccc	891
Glu	Ser	Ser	Arg	Ile	Arg	Asp	Ala	Asn	Arg	Arg	Arg	Pro	Asp	Asp	Pro	
					240					245					250	
ctt	tac	gat	aga	aag	acc	tta	cac	ata	cca	cct	gat	ggt	ttc	aag	aaa	939
Leu	Tyr	Asp	Arg	Lys	Thr	Leu	His	Ile	Pro	Pro	Asp	Val	Phe	Lys	Lys	
				255					260					265		
atg	tct	gca	tca	caa	aag	caa	tat	tgg	agt	ggt	aag	agt	gaa	tat	atg	987
Met	Ser	Ala	Ser	Gln	Lys	Gln	Tyr	Trp	Ser	Val	Lys	Ser	Glu	Tyr	Met	
				270				275					280			
gac	att	gtg	ctt	ttc	ttt	aaa	gtg	ggg	aaa	ttt	tat	gag	ctg	tat	gag	1035
Asp	Ile	Val	Leu	Phe	Phe	Lys	Val	Gly	Lys	Phe	Tyr	Glu	Leu	Tyr	Glu	
			285				290					295				
cta	gat	gcg	gaa	tta	ggg	cac	aag	gag	ctt	gac	tgg	aag	atg	acc	atg	1083
Leu	Asp	Ala	Glu	Leu	Gly	His	Lys	Glu	Leu	Asp	Trp	Lys	Met	Thr	Met	
			300				305				310					
agt	ggg	gtg	gga	aaa	tgc	aga	cag	ggt	ggg	atc	tct	gaa	agt	ggg	ata	1131
Ser	Gly	Val	Gly	Lys	Cys	Arg	Gln	Val	Gly	Ile	Ser	Glu	Ser	Gly	Ile	
					320					325					330	
gat	gag	gca	gtg	caa	aag	cta	tta	gct	cgt	gga	tat	aaa	ggt	gga	cga	1179
Asp	Glu	Ala	Val	Gln	Lys	Leu	Leu	Ala	Arg	Gly	Tyr	Lys	Val	Gly	Arg	
				335					340					345		

atc gag cag cta gaa aca tct gac caa gca aaa gcc aga ggt gct aat	1227
Ile Glu Gln Leu Glu Thr Ser Asp Gln Ala Lys Ala Arg Gly Ala Asn	
350 355 360	
act ata att cca agg aag cta gtt cag gta tta act cca tca aca gca	1275
Thr Ile Ile Pro Arg Lys Leu Val Gln Val Leu Thr Pro Ser Thr Ala	
365 370 375	
agc gag gga aac atc ggg cct gat gcc gtc cat ctt ctt gct ata aaa	1323
Ser Glu Gly Asn Ile Gly Pro Asp Ala Val His Leu Leu Ala Ile Lys	
380 385 390	
gag atc aaa atg gag cta caa aag tgt tca act gtg tat gga ttt gct	1371
Glu Ile Lys Met Glu Leu Gln Lys Cys Ser Thr Val Tyr Gly Phe Ala	
395 400 405 410	
ttt gtt gac tgt gct gcc ttg agg ttt tgg gtt ggg tcc atc agc gat	1419
Phe Val Asp Cys Ala Ala Leu Arg Phe Trp Val Gly Ser Ile Ser Asp	
415 420 425	
gat gca tca tgt gct gct ctt gga gcg tta ttg atg cag gtt tct cca	1467
Asp Ala Ser Cys Ala Ala Leu Gly Ala Leu Leu Met Gln Val Ser Pro	
430 435 440	
aag gaa gtg tta tat gac agt aaa ggg cta tca aga gaa gca caa aag	1515
Lys Glu Val Leu Tyr Asp Ser Lys Gly Leu Ser Arg Glu Ala Gln Lys	
445 450 455	
gct cta agg aaa tat acg ttg aca ggg tct acg gcg gta cag ttg gct	1563
Ala Leu Arg Lys Tyr Thr Leu Thr Gly Ser Thr Ala Val Gln Leu Ala	
460 465 470	
cca gta cca caa gta atg ggg gat aca gat gct gct gga gtt aga aat	1611
Pro Val Pro Gln Val Met Gly Asp Thr Asp Ala Ala Gly Val Arg Asn	
475 480 485 490	
ata ata gaa tct aac gga tac ttt aaa ggt tct tct gaa tca tgg aac	1659
Ile Ile Glu Ser Asn Gly Tyr Phe Lys Gly Ser Ser Glu Ser Trp Asn	
495 500 505	
tgt gct gtt gat ggt cta aat gaa tgt gat gtt gcc ctt agt gct ctt	1707
Cys Ala Val Asp Gly Leu Asn Glu Cys Asp Val Ala Leu Ser Ala Leu	
510 515 520	
gga gag cta att aat cat ctg tct agg cta aag cta gaa gat gta ctt	1755
Gly Glu Leu Ile Asn His Leu Ser Arg Leu Lys Leu Glu Asp Val Leu	
525 530 535	
aag cat ggg gat att ttt cca tac caa gtt tac agg ggt tgt ctc aga	1803
Lys His Gly Asp Ile Phe Pro Tyr Gln Val Tyr Arg Gly Cys Leu Arg	
540 545 550	
att gat ggc cag acg atg gta aat ctt gag ata ttt aac aat agc tgt	1851
Ile Asp Gly Gln Thr Met Val Asn Leu Glu Ile Phe Asn Asn Ser Cys	
555 560 565 570	
gat ggt ggt cct tca ggg acc ttg tac aaa tat ctt gat aac tgt gtt	1899
Asp Gly Gly Pro Ser Gly Thr Leu Tyr Lys Tyr Leu Asp Asn Cys Val	
575 580 585	

agt cca act ggt aag cga ctc tta agg aat tgg atc tgc cat cca ctc	1947
Ser Pro Thr Gly Lys Arg Leu Leu Arg Asn Trp Ile Cys His Pro Leu	
590 595 600	
aaa gat gta gaa agc atc aat aaa cgg ctt gat gta gtt gaa gaa ttc	1995
Lys Asp Val Glu Ser Ile Asn Lys Arg Leu Asp Val Val Glu Glu Phe	
605 610 615	
acg gca aac tca gaa agt atg caa atc act ggc cag tat ctc cac aaa	2043
Thr Ala Asn Ser Glu Ser Met Gln Ile Thr Gly Gln Tyr Leu His Lys	
620 625 630	
ctt cca gac tta gaa aga ctg ctc gga cgc atc aag tct agc gtt cga	2091
Leu Pro Asp Leu Glu Arg Leu Leu Gly Arg Ile Lys Ser Ser Val Arg	
635 640 645 650	
tca tca gcc tct gtg ttg cct gct ctt ctg ggg aaa aaa gtg ctg aaa	2139
Ser Ser Ala Ser Val Leu Pro Ala Leu Leu Gly Lys Lys Val Leu Lys	
655 660 665	
caa cga gtt aaa gca ttt ggg caa att gtg aaa ggg ttc aga agt gga	2187
Gln Arg Val Lys Ala Phe Gly Gln Ile Val Lys Gly Phe Arg Ser Gly	
670 675 680	
att gat ctg ttg ttg gct cta cag aag gaa tca aat atg atg agt ttg	2235
Ile Asp Leu Leu Leu Ala Leu Gln Lys Glu Ser Asn Met Met Ser Leu	
685 690 695	
ctt tat aaa ctc tgt aaa ctt cct ata tta gta gga aaa agc ggg cta	2283
Leu Tyr Lys Leu Cys Lys Leu Pro Ile Leu Val Gly Lys Ser Gly Leu	
700 705 710	
gag tta ttt ctt tct caa ttc gaa gca gcc ata gat agc gac ttt cca	2331
Glu Leu Phe Leu Ser Gln Phe Glu Ala Ala Ile Asp Ser Asp Phe Pro	
715 720 725 730	
aat tat cag aac caa gat gtg aca gat gaa aac gct gaa act ctc aca	2379
Asn Tyr Gln Asn Gln Asp Val Thr Asp Glu Asn Ala Glu Thr Leu Thr	
735 740 745	
ata ctt atc gaa ctt ttt atc gaa aga gca act caa tgg tct gag gtc	2427
Ile Leu Ile Glu Leu Phe Ile Glu Arg Ala Thr Gln Trp Ser Glu Val	
750 755 760	
att cac acc ata agc tgc cta gat gtc ctg aga tct ttt gca atc gca	2475
Ile His Thr Ile Ser Cys Leu Asp Val Leu Arg Ser Phe Ala Ile Ala	
765 770 775	
gca agt ctc tct gct gga agc atg gcc agg cct gtt att ttt ccc gaa	2523
Ala Ser Leu Ser Ala Gly Ser Met Ala Arg Pro Val Ile Phe Pro Glu	
780 785 790	
tca gaa gct aca gat cag aat cag aaa aca aaa ggg cca ata ctt aaa	2571
Ser Glu Ala Thr Asp Gln Asn Gln Lys Thr Lys Gly Pro Ile Leu Lys	
795 800 805 810	
atc caa gga cta tgg cat cca ttt gca gtt gca gcc gat ggt caa ttg	2619
Ile Gln Gly Leu Trp His Pro Phe Ala Val Ala Ala Asp Gly Gln Leu	
815 820 825	

cct gtt ccg aat gat ata ctc ctt ggc gag gct aga aga agc agt ggc	2667
Pro Val Pro Asn Asp Ile Leu Leu Gly Glu Ala Arg Arg Ser Ser Gly	
830 835 840	
agc att cat cct cgg tca ttg tta ctg acg gga cca aac atg ggc gga	2715
Ser Ile His Pro Arg Ser Leu Leu Leu Thr Gly Pro Asn Met Gly Gly	
845 850 855	
aaa tca act ctt ctt cgt gca aca tgt ctg gcc gtt atc ttt gcc caa	2763
Lys Ser Thr Leu Leu Arg Ala Thr Cys Leu Ala Val Ile Phe Ala Gln	
860 865 870	
ctt ggc tgc tac gtg ccg tgt gag tct tgc gaa atc tcc ctc gtg gat	2811
Leu Gly Cys Tyr Val Pro Cys Glu Ser Cys Glu Ile Ser Leu Val Asp	
875 880 885 890	
act atc ttc aca agg ctt ggc gca tct gat aga atc atg aca gga gag	2859
Thr Ile Phe Thr Arg Leu Gly Ala Ser Asp Arg Ile Met Thr Gly Glu	
895 900 905	
agt acc ttt ttg gta gaa tgc act gag aca gcg tca gtt ctt cag aat	2907
Ser Thr Phe Leu Val Glu Cys Thr Glu Thr Ala Ser Val Leu Gln Asn	
910 915 920	
gca act cag gat tca cta gta atc ctt gac gaa ctg ggc aga gga act	2955
Ala Thr Gln Asp Ser Leu Val Ile Leu Asp Glu Leu Gly Arg Gly Thr	
925 930 935	
agt act ttc gat gga tac gcc att gca tac tcg gtt ttt cgt cac ctg	3003
Ser Thr Phe Asp Gly Tyr Ala Ile Ala Tyr Ser Val Phe Arg His Leu	
940 945 950	
gta gag aaa gtt caa tgt cgg atg ctc ttt gca aca cat tac cac cct	3051
Val Glu Lys Val Gln Cys Arg Met Leu Phe Ala Thr His Tyr His Pro	
955 960 965 970	
ctc acc aag gaa ttc gcg tct cac cca cgt gtc acc tcg aaa cac atg	3099
Leu Thr Lys Glu Phe Ala Ser His Pro Arg Val Thr Ser Lys His Met	
975 980 985	
gct tgc gca ttc aaa tca aga tct gat tat caa cca cgt ggt tgt gat	3147
Ala Cys Ala Phe Lys Ser Arg Ser Asp Tyr Gln Pro Arg Gly Cys Asp	
990 995 1000	
caa gac cta gtg ttc ttg tac cgt tta acc gag gga gct tgt cct gag	3195
Gln Asp Leu Val Phe Leu Tyr Arg Leu Thr Glu Gly Ala Cys Pro Glu	
1005 1010 1015	
agc tac gga ctt caa gtg gca ctc atg gct gga ata cca aac caa gtg	3243
Ser Tyr Gly Leu Gln Val Ala Leu Met Ala Gly Ile Pro Asn Gln Val	
1020 1025 1030	
gtt gaa aca gca tca ggt gct gct caa gcc atg aag aga tca att ggg	3291
Val Glu Thr Ala Ser Gly Ala Ala Gln Ala Met Lys Arg Ser Ile Gly	
1035 1040 1045 1050	
gga aac ttc aag tca agt gag cta aga tct gag ttc tca agt ctg cat	3339
Glu Asn Phe Lys Ser Ser Glu Leu Arg Ser Glu Phe Ser Ser Leu His	
1055 1060 1065	

gaa gac tgg ctc aag tca ttg gtg ggt att tct cga gtc gcc cac aac 3387  
 Glu Asp Trp Leu Lys Ser Leu Val Gly Ile Ser Arg Val Ala His Asn  
                   1070                  1075                  1080

aat gcc ccc att ggc gaa gat gac tac gac act ttg ttt tgc tta tgg 3435  
 Asn Ala Pro Ile Gly Glu Asp Asp Tyr Asp Thr Leu Phe Cys Leu Trp  
                   1085                  1090                  1095

cat gag atc aaa tcc tct tac tgt gtt ccc aaa taaatggcta 3478  
 His Glu Ile Lys Ser Ser Tyr Cys Val Pro Lys  
                   1100                  1105

tgacataaca ctatctgaag ctcgttaagt cttttgcctc tctgatgttt attcctctta 3538  
 aaaaatgctt atatatcaaa aaattgtttc ctcgattaaa aaaaaaaaaa aaaaaaaaaa 3598  
 aaaaaaaaaa 3606

<210> 31  
 <211> 1109  
 <212> PRT  
 <213> Arabidopsis thaliana ecotype Columbia  
 <223> Polypeptide MSH6

<400> 31

Met Gln Arg Gln Arg Ser Ile Leu Ser Phe Phe Gln Lys Pro Thr Ala  
   1                  5                  10                  15

Ala Thr Thr Lys Gly Leu Val Ser Gly Asp Ala Ala Ser Gly Gly Gly  
                   20                  25                  30

Gly Ser Gly Gly Pro Arg Phe Asn Val Arg Glu Gly Asp Ala Lys Gly  
                   35                  40                  45

Asp Ala Ser Val Arg Phe Ala Val Ser Lys Ser Val Asp Glu Val Arg  
   50                  55                  60

Gly Thr Asp Thr Pro Pro Glu Lys Val Pro Arg Arg Val Leu Pro Ser  
   65                  70                  75                  80

Gly Phe Lys Pro Ala Glu Ser Ala Gly Asp Ala Ser Ser Leu Phe Ser  
                   85                  90                  95

Asn Ile Met His Lys Phe Val Lys Val Asp Asp Arg Asp Cys Ser Gly  
                   100                  105                  110

Glu Arg Ser Arg Glu Asp Val Val Pro Leu Asn Asp Ser Ser Leu Cys  
                   115                  120                  125

Met Lys Ala Asn Asp Val Ile Pro Gln Phe Arg Ser Asn Asn Gly Lys  
   130                  135                  140

Thr Gln Glu Arg Asn His Ala Phe Ser Phe Ser Gly Arg Ala Glu Leu  
   145                  150                  155                  160

Arg Ser Val Glu Asp Ile Gly Val Asp Gly Asp Val Pro Gly Pro Glu  
                   165                  170                  175

Thr Pro Gly Met Arg Pro Arg Ala Ser Arg Leu Lys Arg Val Leu Glu  
                   180                  185                  190

Asp	Glu	Met	Thr	Phe	Lys	Glu	Asp	Lys	Val	Pro	Val	Leu	Asp	Ser	Asn		
	195						200					205					
Lys	Arg	Leu	Lys	Met	Leu	Gln	Asp	Pro	Val	Cys	Gly	Glu	Lys	Lys	Glu		
	210					215					220						
Val	Asn	Glu	Gly	Thr	Lys	Phe	Glu	Trp	Leu	Glu	Ser	Ser	Arg	Ile	Arg		
225					230					235					240		
Asp	Ala	Asn	Arg	Arg	Arg	Pro	Asp	Asp	Pro	Leu	Tyr	Asp	Arg	Lys	Thr		
			245						250					255			
Leu	His	Ile	Pro	Pro	Asp	Val	Phe	Lys	Lys	Met	Ser	Ala	Ser	Gln	Lys		
		260						265						270			
Gln	Tyr	Trp	Ser	Val	Lys	Ser	Glu	Tyr	Met	Asp	Ile	Val	Leu	Phe	Phe		
	275						280					285					
Lys	Val	Gly	Lys	Phe	Tyr	Glu	Leu	Tyr	Glu	Leu	Asp	Ala	Glu	Leu	Gly		
	290					295					300						
His	Lys	Glu	Leu	Asp	Trp	Lys	Met	Thr	Met	Ser	Gly	Val	Gly	Lys	Cys		
305					310					315					320		
Arg	Gln	Val	Gly	Ile	Ser	Glu	Ser	Gly	Ile	Asp	Glu	Ala	Val	Gln	Lys		
			325						330					335			
Leu	Leu	Ala	Arg	Gly	Tyr	Lys	Val	Gly	Arg	Ile	Glu	Gln	Leu	Glu	Thr		
		340						345					350				
Ser	Asp	Gln	Ala	Lys	Ala	Arg	Gly	Ala	Asn	Thr	Ile	Ile	Pro	Arg	Lys		
	355						360					365					
Leu	Val	Gln	Val	Leu	Thr	Pro	Ser	Thr	Ala	Ser	Glu	Gly	Asn	Ile	Gly		
	370					375					380						
Pro	Asp	Ala	Val	His	Leu	Leu	Ala	Ile	Lys	Glu	Ile	Lys	Met	Glu	Leu		
385					390					395					400		
Gln	Lys	Cys	Ser	Thr	Val	Tyr	Gly	Phe	Ala	Phe	Val	Asp	Cys	Ala	Ala		
			405					410					415				
Leu	Arg	Phe	Trp	Val	Gly	Ser	Ile	Ser	Asp	Asp	Ala	Ser	Cys	Ala	Ala		
	420						425						430				
Leu	Gly	Ala	Leu	Leu	Met	Gln	Val	Ser	Pro	Lys	Glu	Val	Leu	Tyr	Asp		
	435					440						445					
Ser	Lys	Gly	Leu	Ser	Arg	Glu	Ala	Gln	Lys	Ala	Leu	Arg	Lys	Tyr	Thr		
	450					455					460						
Leu	Thr	Gly	Ser	Thr	Ala	Val	Gln	Leu	Ala	Pro	Val	Pro	Gln	Val	Met		
465					470					475					480		
Gly	Asp	Thr	Asp	Ala	Ala	Gly	Val	Arg	Asn	Ile	Ile	Glu	Ser	Asn	Gly		
			485						490					495			
Tyr	Phe	Lys	Gly	Ser	Ser	Glu	Ser	Trp	Asn	Cys	Ala	Val	Asp	Gly	Leu		
			500					505					510				



Asn Glu Cys Asp Val Ala Leu Ser Ala Leu Gly Glu Leu Ile Asn His  
 515 520 525  
 Leu Ser Arg Leu Lys Leu Glu Asp Val Leu Lys His Gly Asp Ile Phe  
 530 535 540  
 Pro Tyr Gln Val Tyr Arg Gly Cys Leu Arg Ile Asp Gly Gln Thr Met  
 545 550 555 560  
 Val Asn Leu Glu Ile Phe Asn Asn Ser Cys Asp Gly Gly Pro Ser Gly  
 565 570 575  
 Thr Leu Tyr Lys Tyr Leu Asp Asn Cys Val Ser Pro Thr Gly Lys Arg  
 580 585 590  
 Leu Leu Arg Asn Trp Ile Cys His Pro Leu Lys Asp Val Glu Ser Ile  
 595 600 605  
 Asn Lys Arg Leu Asp Val Val Glu Glu Phe Thr Ala Asn Ser Glu Ser  
 610 615 620  
 Met Gln Ile Thr Gly Gln Tyr Leu His Lys Leu Pro Asp Leu Glu Arg  
 625 630 635 640  
 Leu Leu Gly Arg Ile Lys Ser Ser Val Arg Ser Ser Ala Ser Val Leu  
 645 650 655  
 Pro Ala Leu Leu Gly Lys Lys Val Leu Lys Gln Arg Val Lys Ala Phe  
 660 665 670  
 Gly Gln Ile Val Lys Gly Phe Arg Ser Gly Ile Asp Leu Leu Leu Ala  
 675 680 685  
 Leu Gln Lys Glu Ser Asn Met Met Ser Leu Leu Tyr Lys Leu Cys Lys  
 690 695 700  
 Leu Pro Ile Leu Val Gly Lys Ser Gly Leu Glu Leu Phe Leu Ser Gln  
 705 710 715 720  
 Phe Glu Ala Ala Ile Asp Ser Asp Phe Pro Asn Tyr Gln Asn Gln Asp  
 725 730 735  
 Val Thr Asp Glu Asn Ala Glu Thr Leu Thr Ile Leu Ile Glu Leu Phe  
 740 745 750  
 Ile Glu Arg Ala Thr Gln Trp Ser Glu Val Ile His Thr Ile Ser Cys  
 755 760 765  
 Leu Asp Val Leu Arg Ser Phe Ala Ile Ala Ala Ser Leu Ser Ala Gly  
 770 775 780  
 Ser Met Ala Arg Pro Val Ile Phe Pro Glu Ser Glu Ala Thr Asp Gln  
 785 790 795 800  
 Asn Gln Lys Thr Lys Gly Pro Ile Leu Lys Ile Gln Gly Leu Trp His  
 805 810 815  
 Pro Phe Ala Val Ala Ala Asp Gly Gln Leu Pro Val Pro Asn Asp Ile  
 820 825 830

Leu Leu Gly Glu Ala Arg Arg Ser Ser Gly Ser Ile His Pro Arg Ser  
           835                                  840                                  845  
 Leu Leu Leu Thr Gly Pro Asn Met Gly Gly Lys Ser Thr Leu Leu Arg  
           850                                  855                                  860  
 Ala Thr Cys Leu Ala Val Ile Phe Ala Gln Leu Gly Cys Tyr Val Pro  
           865                                  870                                  875                                  880  
 Cys Glu Ser Cys Glu Ile Ser Leu Val Asp Thr Ile Phe Thr Arg Leu  
                                   885                                  890                                  895  
 Gly Ala Ser Asp Arg Ile Met Thr Gly Glu Ser Thr Phe Leu Val Glu  
                                   900                                  905                                  910  
 Cys Thr Glu Thr Ala Ser Val Leu Gln Asn Ala Thr Gln Asp Ser Leu  
                                   915                                  920                                  925  
 Val Ile Leu Asp Glu Leu Gly Arg Gly Thr Ser Thr Phe Asp Gly Tyr  
           930                                  935                                  940  
 Ala Ile Ala Tyr Ser Val Phe Arg His Leu Val Glu Lys Val Gln Cys  
           945                                  950                                  955                                  960  
 Arg Met Leu Phe Ala Thr His Tyr His Pro Leu Thr Lys Glu Phe Ala  
                                   965                                  970                                  975  
 Ser His Pro Arg Val Thr Ser Lys His Met Ala Cys Ala Phe Lys Ser  
                                   980                                  985                                  990  
 Arg Ser Asp Tyr Gln Pro Arg Gly Cys Asp Gln Asp Leu Val Phe Leu  
                                   995                                  1000                                  1005  
 Tyr Arg Leu Thr Glu Gly Ala Cys Pro Glu Ser Tyr Gly Leu Gln Val  
           1010                                  1015                                  1020  
 Ala Leu Met Ala Gly Ile Pro Asn Gln Val Val Glu Thr Ala Ser Gly  
           1025                                  1030                                  1035                                  1040  
 Ala Ala Gln Ala Met Lys Arg Ser Ile Gly Glu Asn Phe Lys Ser Ser  
                                   1045                                  1050                                  1055  
 Glu Leu Arg Ser Glu Phe Ser Ser Leu His Glu Asp Trp Leu Lys Ser  
                                   1060                                  1065                                  1070  
 Leu Val Gly Ile Ser Arg Val Ala His Asn Asn Ala Pro Ile Gly Glu  
           1075                                  1080                                  1085  
 Asp Asp Tyr Asp Thr Leu Phe Cys Leu Trp His Glu Ile Lys Ser Ser  
           1090                                  1095                                  1100  
 Tyr Cys Val Pro Lys  
 1105

<210> 32  
 <211> 24  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> Forward primer for PCR amplification of ATHGENEA microsatellite  
 <400> 32  
 accatgcata gcttaaactt cttg 24

<210> 33  
 <211> 22  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> Reverse primer for PCR amplification of ATHGENEA microsatellite  
 <400> 33  
 acataaccac aaataggggt gc 22

<210> 34  
 <211> 18  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> Forward primer DMCIN-A for PCR on genomic DNA of Arabidopsis thaliana  
 ssp. Landsberg erecta "Ler"  
 <400> 34  
 gaagcgatat tgttcgtg 18

<210> 35  
 <211> 18  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> Reverse primer DMCIN-B for PCR on genomic DNA of Arabidopsis thaliana  
 ssp. Landsberg erecta "Ler"  
 <400> 35  
 agattgcgag aacattcc 18

<210> 36  
 <211> 31  
 <212> DNA  
 <213> Artificial sequence

<220>  
 <223> Forward primer DMCIN-1 for PCR on genomic DNA of Arabidopsis thaliana  
 ssp. Landsberg erecta "Ler"  
 <400> 36

acgcgtcgac tcagctatga gattactcgt g 31

<210> 37  
<211> 29  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Reverse primer DMCIN-2 for PCR on genomic DNA of Arabidopsis thaliana  
ssp. Landsberg erecta "Ler"

<400> 37

gctctagatt tctcgctcta agactctct 29

<210> 38  
<211> 32  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Forward primer DMCIN-3 for PCR on genomic DNA of Arabidopsis thaliana  
ssp. Landsberg erecta "Ler"

<400> 38

gctctagagc ttctcttaag taagtgattg at 32

<210> 39  
<211> 48  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Reverse primer DMCIN-4 for PCR on genomic DNA of Arabidopsis thaliana  
ssp. Landsberg erecta "Ler"

<400> 39

tccccgggc tcgagagatc tccatggttt cttcagctct atgaatcc 48

<210> 40  
<211> 26  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Forward primer DMC1a for PCR on genomic DNA of Arabidopsis thaliana  
ssp. Landsberg erecta "Ler"

<400> 40

acgcgtcgac gaattcgcaa gtgggg 26

<210> 41  
 <211> 38  
 <212> DNA  
 <213> Artificial sequence  
  
 <220>  
 <223> Reverse primer DMC1b for PCR on genomic DNA of *Arabidopsis thaliana* ssp. Landsberg erecta "Ler"  
  
 <400> 41  
 tccatggaga tctcccgggt accgatttgc ttcgaggg 38  
  
 <210> 42  
 <211> 20  
 <212> DNA  
 <213> Artificial sequence  
  
 <220>  
 <223> Forward primer for PCR amplification of ATEAT1 SSLP marker in *Arabidopsis thaliana* subspecies  
  
 <400> 42  
 gccactgcgt gaatgatatg 20  
  
 <210> 43  
 <211> 22  
 <212> DNA  
 <213> Artificial sequence  
  
 <220>  
 <223> Reverse primer for PCR amplification of ATEAT1 SSLP marker in *Arabidopsis thaliana* subspecies  
  
 <400> 43  
 cgaacagcca acattaattc cc 22  
  
 <210> 44  
 <211> 18  
 <212> DNA  
 <213> Artificial sequence  
  
 <220>  
 <223> Forward primer for PCR amplification of NGA63 SSLP marker in *Arabidopsis thaliana* subspecies  
  
 <400> 44  
 aaccaaggca cagaagcg 18  
  
 <210> 45  
 <211> 18

<212> DNA  
<213> Artificial sequence

<220>  
<223> Reverse primer for PCR amplification of NGA63 SSLP marker in  
Arabidopsis  
thaliana subspecies

<400> 45

acccaagtga tcgccacc 18

<210> 46  
<211> 21  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Forward primer for PCR amplification of NGA248 SSLP marker in  
Arabidopsis thaliana subspecies

<400> 46

taccgaacca aaacacaaaag g 21

<210> 47  
<211> 22  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Reverse primer for PCR amplification of NGA248 SSLP marker in  
Arabidopsis thaliana subspecies

<400> 47

tctgtatctc ggtgaattct cc 22

<210> 48  
<211> 22  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Forward primer for PCR amplification of NGA128 SSLP marker in  
Arabidopsis thaliana subspecies

<400> 48

ggtctgttga tgtcgtaagt cg 22

<210> 49  
<211> 22  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Reverse primer for PCR amplification of NGA128 SSLP marker in  
Arabidopsis thaliana subspecies

<400> 49

atcttgaaac ctttagggag gg 22

<210> 50  
<211> 22  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Forward primer for PCR amplification of NGA280 SSLP marker in  
Arabidopsis thaliana subspecies

<400> 50

ctgatctcac ggacaatagt gc 22

<210> 51  
<211> 20  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Reverse primer for PCR amplification of NGA280 SSLP marker in  
Arabidopsis thaliana subspecies

<400> 51

ggctccataa aaagtcacc 20

<210> 52  
<211> 21  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Forward primer for PCR amplification of NGA111 SSLP marker in  
Arabidopsis thaliana subspecies

<400> 52

ctccagttgg aagctaaagg g 21

<210> 53  
<211> 21  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Reverse primer for PCR amplification of NGA111 SSLP marker in  
Arabidopsis thaliana subspecies

<400> 53

tgtttttttag gacaaatggc g 21

<210> 54

<211> 20

<212> DNA

<213> Artificial sequence

<220>

<223> Forward primer for PCR amplification of NGA168 SSLP marker in  
Arabidopsis thaliana subspecies

<400> 54

ccttcacatc caaaaccac 20

<210> 55

<211> 20

<212> DNA

<213> Artificial sequence

<220>

<223> Reverse primer for PCR amplification of NGA168 SSLP marker in  
Arabidopsis thaliana subspecies

<400> 55

gcacataccc acaaccagaa 20

<210> 56

<211> 20

<212> DNA

<213> Artificial sequence

<220>

<223> Forward primer for PCR amplification of NGA1126 SSLP marker in  
Arabidopsis thaliana subspecies

<400> 56

cgctacgctt ttcggtaaag 20

<210> 57

<211> 20

<212> DNA

<213> Artificial sequence

<220>

<223> Reverse primer for PCR amplification of NGA1126 SSLP marker in  
Arabidopsis thaliana subspecies

<400> 57

gcacagtcca agtcacaacc 20



<210> 58  
<211> 20  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Forward primer for PCR amplification of NGA361 SSLP marker in  
Arabidopsis thaliana subspecies

<400> 58

aaagagatga gaatttggac 20

<210> 59  
<211> 23  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Reverse primer for PCR amplification of NGA361 SSLP marker in  
Arabidopsis thaliana subspecies

<400> 59

acatatcaat atattaaagt agc 23

<210> 60  
<211> 18  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Forward primer for PCR amplification of NGA168 SSLP marker in  
Arabidopsis thaliana subspecies

<400> 60

tcgtctactg cactgccg 18

<210> 61  
<211> 22  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Reverse primer for PCR amplification of NGA168 SSLP marker in  
Arabidopsis thaliana subspecies

<400> 61

gaggacatgt ataggagcct cg 22

<210> 62  
<211> 20  
<212> DNA  
<213> Artificial sequence

<220>  
 <223> Forward primer for PCR amplification of AthBIO2 SSLP marker in  
 Arabidopsis thaliana subspecies  
  
 <400> 62  
 tgacctcctc ttccatggag 20  
  
 <210> 63  
 <211> 22  
 <212> DNA  
 <213> Artificial sequence  
  
 <220>  
 <223> Reverse primer for PCR amplification of AthBIO2 SSLP marker in  
 Arabidopsis thaliana subspecies  
  
 <400> 63  
 ttaacagaaa cccaaagctt tc 22  
  
 <210> 64  
 <211> 21  
 <212> DNA  
 <213> Artificial sequence  
  
 <220>  
 <223> Forward primer for PCR amplification of AthUBIQUE SSLP marker in  
 Arabidopsis thaliana subspecies  
  
 <400> 64  
 aggcaaatgt ccatttcatt g 21  
  
 <210> 65  
 <211> 20  
 <212> DNA  
 <213> Artificial sequence  
  
 <220>  
 <223> Reverse primer for PCR amplification of AthUBIQUE SSLP marker in  
 Arabidopsis thaliana subspecies  
  
 <400> 65  
 acgacatggc agattttctcc 20  
  
 <210> 66  
 <211> 21  
 <212> DNA  
 <213> Artificial sequence  
  
 <220>  
 <223> Forward primer for PCR amplification of NGA172 SSLP marker in  
 Arabidopsis thaliana subspecies

<400> 66

agctgcttcc ttatagcgtc c 21

<210> 67

<211> 19

<212> DNA

<213> Artificial sequence

<220>

<223> Reverse primer for PCR amplification of NGA172 SSLP marker in *Arabidopsis thaliana* subspecies

<400> 67

catccgaatg ccattgttc 19

<210> 68

<211> 21

<212> DNA

<213> Artificial sequence

<220>

<223> Forward primer for PCR amplification of NGA126 SSLP marker in *Arabidopsis thaliana* subspecies

<400> 68

gaaaaaacgc tactttcgtg g 21

<210> 69

<211> 22

<212> DNA

<213> Artificial sequence

<220>

<223> Reverse primer for PCR amplification of NGA126 SSLP marker in *Arabidopsis thaliana* subspecies

<400> 69

caagagcaat atcaagagca gc 22

<210> 70

<211> 20

<212> DNA

<213> Artificial sequence

<220>

<223> Forward primer for PCR amplification of NGA162 SSLP marker in *Arabidopsis thaliana* subspecies

<400> 70

catgcaattt gcatctgagg 20

<210> 71  
<211> 22  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Reverse primer for PCR amplification of NGA162 SSLP marker in  
Arabidopsis thaliana subspecies

<400> 71

ctctgtcact cttttcctct gg 22

<210> 72  
<211> 21  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Forward primer for PCR amplification of NGA6 SSLP marker in Arabidopsis  
thaliana subspecies

<400> 72

tggatttctt cctctcttca c 21

<210> 73  
<211> 21  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Reverse primer for PCR amplification of NGA6 SSLP marker in Arabidopsis  
thaliana subspecies

<400> 73

atggagaagc ttacactgat c 21

<210> 74  
<211> 20  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Forward primer for PCR amplification of NGA12 SSLP marker in  
Arabidopsis  
thaliana subspecies

<400> 74

aatgttgccc tcccctcctc 20

<210> 75  
<211> 22  
<212> DNA

<213> Artificial sequence

<220>

<223> Reverse primer for PCR amplification of NGA12 SSLP marker in Arabidopsis thaliana subspecies

<400> 75

tgatgctctc tgaaacaaga gc 22

<210> 76

<211> 21

<212> DNA

<213> Artificial sequence

<220>

<223> Forward primer for PCR amplification of NGA8 SSLP marker in Arabidopsis thaliana subspecies

<400> 76

gagggcaaat ctttatttcg g 21

<210> 77

<211> 22

<212> DNA

<213> Artificial sequence

<220>

<223> Reverse primer for PCR amplification of NGA8 SSLP marker in Arabidopsis thaliana subspecies

<400> 77

tggctttcgt ttataaacat cc 22

<210> 78

<211> 21

<212> DNA

<213> Artificial sequence

<220>

<223> Forward primer for PCR amplification of NGA1107 SSLP marker in Arabidopsis thaliana subspecies

<400> 78

gcgaaaaaac aaaaaaatcc a 21

<210> 79

<211> 21

<212> DNA

<213> Artificial sequence

<220>

<223> Reverse primer for PCR amplification of NGA1107 SSLP marker in  
*Arabidopsis thaliana* subspecies

<400> 79

cgacgaatcg acagaattag g 21

<210> 80

<211> 21

<212> DNA

<213> Artificial sequence

<220>

<223> Forward primer for PCR amplification of NGA225 SSLP marker in  
*Arabidopsis thaliana* subspecies

<400> 80

gaaatccaaa tcccagagag g 21

<210> 81

<211> 22

<212> DNA

<213> Artificial sequence

<220>

<223> Reverse primer for PCR amplification of NGA225 SSLP marker in  
*Arabidopsis thaliana* subspecies

<400> 81

tctccccact agtttttgtgt cc 22

<210> 82

<211> 19

<212> DNA

<213> Artificial sequence

<220>

<223> Forward primer for PCR amplification of NGA249 SSLP marker in  
*Arabidopsis thaliana* subspecies

<400> 82

taccgtcaat ttcacgcc 19

<210> 83

<211> 22

<212> DNA

<213> Artificial sequence

<220>

<223> Reverse primer for PCR amplification of NGA249 SSLP marker in  
*Arabidopsis thaliana* subspecies

<400> 83

ggatccctaa ctgtaaaatc cc 22

<210> 84  
<211> 22  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Forward primer for PCR amplification of CA72 SSLP marker in Arabidopsis thaliana subspecies

<400> 84

aatcccagta accaaacaca ca 22

<210> 85  
<211> 20  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Reverse primer for PCR amplification of CA72 SSLP marker in Arabidopsis thaliana subspecies

<400> 85

cccagtctaa ccacgaccac 20

<210> 86  
<211> 20  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Forward primer for PCR amplification of NGA151 SSLP marker in Arabidopsis thaliana subspecies

<400> 86

gttttgggaa gttttgctgg 20

<210> 87  
<211> 24  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Reverse primer for PCR amplification of NGA151 SSLP marker in Arabidopsis thaliana subspecies

<400> 87

cagtctaaaa gcgagagtat gatg 24

<210> 88

<211> 22  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Forward primer for PCR amplification of NGA106 SSLP marker in  
Arabidopsis thaliana subspecies

<400> 88

gttatggagt ttctagggca cg 22

<210> 89  
<211> 20  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Reverse primer for PCR amplification of NGA106 SSLP marker in  
Arabidopsis thaliana subspecies

<400> 89

tgccccattt tgttcttctc 20

<210> 90  
<211> 20  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Forward primer for PCR amplification of NGA139 SSLP marker in  
Arabidopsis thaliana subspecies

<400> 90

agagctacca gatccgatgg 20

<210> 91  
<211> 21  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Reverse primer for PCR amplification of NGA139 SSLP marker in  
Arabidopsis thaliana subspecies

<400> 91

ggtttcgttt cactatccag g 21

<210> 92  
<211> 22  
<212> DNA  
<213> Artificial sequence



<220>  
<223> Forward primer for PCR amplification of NGA76 SSLP marker in  
Arabidopsis  
thaliana subspecies

<400> 92

ggagaaaatg tcactctcca cc 22

<210> 93  
<211> 20  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Reverse primer for PCR amplification of NGA76 SSLP marker in  
Arabidopsis  
thaliana subspecies

<400> 93

aggcatggga gacatttacg 20

<210> 94  
<211> 20  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Forward primer for PCR amplification of ATHS0191 SSLP marker in  
Arabidopsis thaliana subspecies

<400> 94

ctccaccaat catgcaaag 20

<210> 95  
<211> 21  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Reverse primer for PCR amplification of ATHS0191 SSLP marker in  
Arabidopsis thaliana subspecies

<400> 95

tgatgttgat ggagatgggc a 21

<210> 96  
<211> 22  
<212> DNA  
<213> Artificial sequence

<220>  
<223> Forward primer for PCR amplification of NGA129 SSLP marker in

# Arabidopsis thaliana subspecies

<400> 96

tcaggaggaa ctaaagtgag gg 22

<210> 97

<211> 22

<212> DNA

<213> Artificial sequence

<220>

<223> Reverse primer for PCR amplification of NGA129 SSLP marker in Arabidopsis thaliana subspecies

<400> 97

cacactgaag atggtcttga gg 22

<210> 98

<211> 8062

<212> DNA

<213> Arabidopsis thaliana ecotype Columbia

<220>

<223> Genomic DNA sequence of AtMSH6

<400> 98

ttttttggtt	gctaacaata	aaggatatac	gttttatgtc	atcaatataa	ctatatataa	60
aagaaatgaa	agatatatat	tgttttttca	tttatcaaac	aaaacaacaa	gacttttttt	120
ttacttttta	cattggtcaa	caaaatataa	gataaacgac	atcgtttaat	catttcccaa	180
ttttaccctt	aagtttaaca	cctagaacct	tctccatctt	cgcaagcaca	gcctgattag	240
gaacagcttt	accattctca	tattcctgaa	ctacctgagt	cctctcattg	atctgtttcg	300
ccaaatccgc	ttgtgacatc	ttcttctcca	atctcgcttt	ctgtatcatc	aacctcacct	360
ctgctttcac	acgatccatc	gccgcaggct	ctgtttcttc	ttccagcttc	ttcgtgttaa	420
tcaccggaac	cgccgtagat	ttcccttttt	tgttcgaacc	ggcatcgaat	ttcttaaccg	480
tttgaaccgc	gacaccgttt	ctcagagctg	cgtaaacgcg	tttcggatcg	cgtaggctct	540
ggctcttttg	ttttgatttg	tggagaacta	ctggttccca	gtcttggtgt	actgctcctg	600
ggtatctgct	cggcatcgtc	gatgaattga	gagaaaggaa	caacgcgaaa	attttattaa	660
tctgagtttt	gaaattgaga	aacgatgaag	atgaagaatg	ttgttgagag	gattgtgata	720
tttatatata	cgaagattgg	tttctggaga	attcgatcat	ctttttctcc	attttcgtct	780
ctggaacgtt	cttagagatg	attgacgacg	tgtcattatc	tgatttgag	ttaaccaatg	840
ctttttgggt	tggattcgct	gtacaccata	ttatccgatt	tggctcaatg	gttttatata	900
aatttggttt	tcggttcggt	tatgagttat	cattaaaatt	aagctaacca	aaaattttcg	960
taaaatttat	ttcggtttca	attcggatcc	cttacttcca	gaaccgaatt	attcgaaacc	1020
ggggttagcc	gaaccgaata	ccaatgcctg	attgactcgt	tggctagaaa	gatccaacgg	1080
tatacaataa	tagaacataa	atcggacggt	catcaaagcc	tcaaagagtg	aacagtcaac	1140
aaaaaaagtt	gagccctgag	gagtatcggt	tccgccatth	ctacgacgca	aggcgaaaaat	1200
ttttggcgcc	aatctttccc	ccctttcgaa	ttctctcagc	tcaaaacatc	gtttctctct	1260
cactctctct	cacaattcca	aaaaatgcag	cgccagagat	cgattttgtc	tttcttccaa	1320
aaaccacacg	cgccgactac	gaagggtttg	gtttccggcg	atgctgctag	cggcgggggc	1380
ggcagcggag	accacgattt	aatgtgaagg	aaggggatgc	taaaggcgac	gcttctgtac	1440
gttttgctgt	ttcgaaatct	gtcgatgagg	ttagaggaaac	ggatactcca	ccggagaagg	1500
ttccgcgtcg	tgctctgccg	tctggattta	agccggctga	atccgcgggt	gatgcttcgt	1560
ccctgttctc	caatattatg	cataagtttg	taaaagtcga	tgatcgagat	tgttctggag	1620
agaggtacta	atcttcgatt	ctcttaattt	tgttatcttt	agctggaaga	agaagattcg	1680
tgtaatttgt	tgtattcgtt	ggagagattc	tgattactgc	attggatcgt	tgtttacaaa	1740

ttttcaggag	ccgagaagat	ggtgttccgc	tgaatgattc	atctctatgt	atgaaggcta	1800
atgatgttat	tcctcaattt	cgttccaata	atggtaaaac	tcaagaaaga	aacctatgctt	1860
ttagttttcag	tgggagagct	gaacttagat	cagtagaaga	tataggagta	gatggcgatg	1920
ttcctgggtcc	agaaacacca	gggatgcgtc	cacgtgcttc	tcgcttgaag	cgagttctgg	1980
aggatgaaat	gacttttaag	gaggataagg	ttcctgtatt	ggactctaac	aaaaggctga	2040
aaatgctcca	ggatccgggt	tgtggagaga	agaaagaagt	aaacgaagga	accaaatttg	2100
aatggcttga	gtcttctcga	atcagggatg	ccaatagaag	acgtcctgat	gatccccctt	2160
acgatagaaa	gaccttacac	ataccacctg	atgttttcaa	gaaaatgtct	gcatacaciaa	2220
agcaatattg	gagtgttaag	agtgaatata	tggacattgt	gcttttcttt	aaagtgggta	2280
gtaactatta	atctagtgtt	caatccattt	cctcaatgtg	atgtgttcac	ttacatctgt	2340
ttacgttatg	ctcttctcag	gggaaatttt	atgagctgta	tgagctagat	gcggaattag	2400
gtcacaagga	gcttgactgg	aagatgacca	tgagtgggtg	gggaaaatgc	agacaggtaa	2460
attagttgaa	acaactggcc	tgcttgaatt	attgtgtcta	taaattttga	caccaccttt	2520
tgtttcagggt	tggatctctc	gaaagtggga	tagatgaggg	agtgcaaaag	ctattagctc	2580
gtgggtaagg	gaaccatcat	actttatgga	attcgtttac	tgctacttcg	gctaggattt	2640
aagaaatgga	aatcacttca	agcatcatta	gttaggatac	tgagaactca	ggatgttttc	2700
ttattcgtta	tataataagt	cttttcatca	aggagtaaca	aacaaaactt	gcacaatatt	2760
tgtgtgtctca	ctggcaaggc	atatataccc	agctaaccct	tgctagtcca	ctgtagtaac	2820
agttacggat	aatatatgtt	tacttgtatg	tggtaccctc	atgttgtctc	tcattggaggc	2880
tttcaagcct	tgtgttgaaa	ctggatagtt	acatatgctt	ccaacagaaa	ctagcatgca	2940
gattcatatg	ctttcctatt	ctactaatta	tgtattgaca	cactcgttgt	ttcttttgaa	3000
agatataaag	ttggacgaat	cgagcagcta	gaaacatctg	accaagcaaa	agccagagggt	3060
gctaatactg	taagttttct	tggatagggt	aaggagagtg	ttgcagactg	tttttgatca	3120
tttctttttc	tgtacattac	tttcatgctg	taattaactc	aatggctatt	ctgggtctgat	3180
tatcagataa	ttccaaggaa	gctagtccag	gtattaaact	catcaacagc	aagcggaggga	3240
aacatcgggc	ctgatgccgt	ccatcttctt	gctataaaaag	aggtttggtt	tttacttatt	3300
tatcttatca	tgttcagttc	atccaagctc	tgaaaaatta	cactcttctt	taccaatctt	3360
ccatcaagct	gtgtaaaagg	tttggaatta	gaaaatcatt	atgtgatgct	ttgttttata	3420
tgcaagagggt	tcccttgaaa	agatctgttt	aagattcttt	gcacttgaaa	aattcaatct	3480
ttttaagtga	atccccctact	ttcttacaat	gatcatagtc	tgcaattgca	tgtcaagtaa	3540
tatcatctct	tgttactgca	tccccctctt	tcttaatgac	cattgtctat	gttgtgtttg	3600
tctcgtgtgc	tggagaaagt	gatagctgat	ccaagtgtga	cattatcatg	attaagtagc	3660
tgctcaggaa	ttgcctttgg	ttacattgcc	taatggtttg	atgtcaattt	ttcttctgaa	3720
tctttatttt	agatcaaaaat	ggagctacaa	aagtgttcaa	ctgtgtatgg	atgtgctttt	3780
gttgactgtg	ctgccttgag	gttttggtt	gggtccatca	gcgatgatgc	atcatgtgct	3840
gctcctggag	cgttattgat	gcaggtaagc	aagtgtattc	tgtatcttat	gtgtaccatg	3900
tgacttcctg	tgcataatatt	tgggttgag	gaactaattc	tgaatcacca	tttgggtatg	3960
tttttccagg	tttctccaaa	ggaagtgtta	tatgacagta	aaggtaaaact	gcttgtatcg	4020
ccagttgttt	tgttaaacag	aatttaagggt	aaatgacact	gggttaattta	aagtgcatac	4080
atgttgtaaat	attgcagggc	tatcaagaga	agcacaaaag	gctctaagga	aatatacggt	4140
gacagggtacc	atttcagtag	gcaagctaac	tgacaattta	accgctcacc	gaatgatagg	4200
tctcttaaac	attgctaattg	tagatgatgt	ttatgtttca	atctaatagg	gtctacggcg	4260
gtacagttgg	ctccagtagc	acaagtaatg	ggggatacag	atgctgctgg	agttagaaat	4320
ataatagaat	ctaaccggata	ctttaaagggt	tcttctgaat	catggaactg	tgtgtgtgat	4380
ggctctaaatg	aatgtgatgt	tgcctttagt	gctcttggag	agctaattaa	tcactctgtct	4440
aggctaaaag	tgtgttgggt	tgtttagttt	ttgcttttca	caaattaagc	aaagggaactt	4500
ttcataactt	acagtttcta	tctacttgca	gctagaagat	gtacttaagc	atggggatat	4560
ttttccatac	caagttttaca	gggggttgtct	cagaattgat	ggccagacga	tggtaaatct	4620
tgagatattt	aacaatagct	gtgatgggtg	tccttcaggc	aagtgcatac	ttcttttttg	4680
ataacttcaa	ctagagggca	gacatagaag	gaaaaattct	aatacttcgt	acggatctcc	4740
agtaagtaat	agccgatttt	tgtttaccta	tgtagggacc	ttgtacaaat	atcttgataa	4800
ctgtgttagt	ccaactggta	agcgactctt	aagggaattg	atctgccatc	cactcaaaga	4860
tgtagaaagc	atcaataaac	ggcttgatgt	agttgaaaga	ttcacggcaa	actcagaaag	4920
tatgcaaatc	actggccagt	atctccacaa	acttccagac	ttagaaaagac	tgctcggacg	4980
catcaagtct	agcgttcgat	catcagcctc	tgtgttgctt	gctcttctgg	ggaaaaaagt	5040
gctgaaacaa	cgagtaagta	tcaatcacaa	gttttctgag	taatgccttc	catgagtagt	5100
ataggactaa	aacattacgg	gtctagctaa	agactgttct	ccttcttttg	caatgtctgg	5160
ttattcatta	catttctctt	aacttattgc	attgcagggt	aaagcatttg	ggcaaatgtg	5220
gaaagggttc	agaagtggaa	ttgatctgtt	gttggtctta	cagaagggaat	caaatatgat	5280
gagtttgctt	tataaactct	gtaaaacttcc	tatattagta	ggaaaaagcg	ggctagagtt	5340

atttctttct	caattcgaag	cagccataga	tagcgacttt	ccaaattatc	aggtgcccac	5400
ctatctttca	tacttttaca	caaaatgtct	gtcactactc	aaagcaatgc	atatggctta	5460
gatctcaact	cacaccccg	ggatccctaaa	gggatttgc	ttttattcct	aatgtttttg	5520
gatggtttga	tttattttct	acttgaactt	attaatcttg	taccagaacc	aagatgtgac	5580
agatgaaaac	gctgaaactc	tcacaatact	tatcgaactt	tttatcgaaa	gagcaactca	5640
atgggtctgag	gtcattcaca	ccataagctg	cctagatgtc	ctgagatctt	ttgcaatcgc	5700
agcaagtctc	tctgctggaa	gcatggccag	gcctgttatt	tttcccgaat	cagaagctac	5760
agatcagaat	cagaaaacaa	aagggccaat	acttaaaatc	caaggactat	ggcatccatt	5820
tgcagttgca	gccgatggtc	aattgcctgt	tccgaatgat	atactccttg	gcgaggctag	5880
aagaagcagt	ggcagcattc	atcctcggtc	attgttactg	acgggaccaa	acatgggctg	5940
aaaatcaact	cttcttcgtg	caacatgtct	ggcctgtatc	tttgcccaag	tttgataact	6000
cgttagataa	ttactctatt	ctttgcaatc	agttcttcaa	catgaataat	aaattctgtt	6060
ttctgtctgc	agcttggtcg	ctacgtgccg	tgtgagtctt	gcgaaatctc	cctcgtggat	6120
actatcttca	caaggccttg	cgcattctgat	agaatcatga	caggagagag	taagttttgt	6180
tctcaaaaata	ccaattcctc	gaactattta	ctcagatttt	gtctgattgg	acaagggtgt	6240
tttgcttttt	tttaggtacc	tttttggtag	aatgcactga	gacagcgtca	gttcttcaga	6300
atgtaactca	ggattccacta	gtaatccttg	acgaactggg	cagaggaact	agtactttcg	6360
atggatacgc	cattgcatac	tcggtaacct	gctcttctcc	ttcaacttat	acttggtgat	6420
caacaaaaac	atgcaattca	ttttgctgaa	acttattgat	ttatatcagg	tttttcgtca	6480
cctggtagag	aaagttcaat	gtcggatgct	ctttgcaaca	cattaccacc	ctctcaccaa	6540
ggaattcgcg	tctcaccac	gtgtcacctc	gaaacacatg	gcttgccgat	tcaaatcaag	6600
atctgattat	caaccacgtg	gttgatgatca	agacctagt	ttcttgatcc	gtttaaccga	6660
gggagcttgt	cctgagagct	acggacttca	agtggcactc	atggctggaa	taccaaacca	6720
agtggttgaa	acagcatcag	gtgctgctca	agccatgaag	agatcaattg	gggaaaactt	6780
caagtcaagt	gagctaagat	ctgagttctc	aagtctgcat	gaagactggc	tcaagtcatt	6840
ggtgggtatt	tctcgagtgc	cccacaacaa	tgccccatt	ggcgaagatg	actacgacac	6900
tttgttttgc	ttatggcatg	agatcaaate	ctcttactgt	gttcccaa	aaatggctat	6960
gacataacac	tatctgaagc	tcgttaagtc	ttttgcttct	ctgatgttta	ttcctcttaa	7020
aaaatgctta	tatatcaaaa	aattgtttcc	tcgattataa	caagattata	tatgtatctg	7080
tcggtttagc	tatggtatat	aatatatgta	tgttcatgag	attggtcaag	agaaatactc	7140
acaaacagta	tattaagaag	gaaatatgtt	tatgcattaa	tttaagtttc	aagataaact	7200
gcaaataacc	tcgactaaag	ttgcaaagac	caaacacaaa	ttacaaaact	tataagactt	7260
aagttctgaa	ttccctaaaa	ccaaaaaaaa	aaacagaaca	tattttgttg	catctacaaa	7320
caacacaaac	ctacatagtt	tataacttac	tcactactga	gattaacatc	agaatcattc	7380
tccatttctt	catcttcaat	ctcatcatca	tcaccaccac	catgatgatt	ctcctcctct	7440
tcacgtaacc	tagcaatctc	actctgagct	ctatcaacaa	tctgcttctt	ctgcaactcc	7500
aaatctctct	gaaaatcagc	tctcatcttc	tccaactcct	tcatttgctc	tttcttactc	7560
ttctccatct	tctcataaac	cttcccaaac	ctctcaacag	aatccgcaa	catcttatac	7620
gaagcagcgt	cattaaacct	cttctctctg	tactcaacct	catcatcctc	atcctcctcc	7680
tcttcagaat	caccaggact	atccatcatc	tcatacaaac	cattagactt	atctaaataa	7740
accttagtgt	tcataaacac	aaactcacct	gaatcaaac	cacaagctaa	acctaaatcc	7800
gacttgggctg	aaacacaaag	caacatatcc	aacttattga	aaaacgacca	tttacttgaa	7860
cctaaacctg	atttctcaac	cttaatcttc	tcttttctat	acttctctct	caagtcacat	7920
atcattctcc	tacattgcgt	ctcagatttc	tccatcctta	gctcctcact	cactttctca	7980
gctacttcat	tccaatcctc	gttctctcaa	ctccttctac	ccaattgcaa	aaacctatct	8040
ccccaaactt	caagcaacac	aa			8062	

B)  
cmd